(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
13 September 2001 (13.09.2001)

PCT

(10) International Publication Number
WO 01/66594 A2

(51) International Patent Classification[7]: C07K 14/47

(21) International Application Number: PCT/US01/06838

(22) International Filing Date: 2 March 2001 (02.03.2001)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/187,150	6 March 2000 (06.03.2000)	US
60/193,404	29 March 2000 (29.03.2000)	US
60/247,013	13 November 2000 (13.11.2000)	US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US	60/187,150 (CIP)
Filed on	6 March 2000 (06.03.2000)
US	60/193,404 (CIP)
Filed on	29 March 2000 (29.03.2000)
US	60/247,013 (CIP)
Filed on	13 November 2000 (13.11.2000)

(71) Applicant *(for all designated States except US)*: **SUGEN, INC.** [US/US]; 230 East Grand Avenue, South San Francisco, CA-94080-4811 (US).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: **PLOWMAN, Gregory, D.** [US/US]; 35 Winding Way, San Carlos, CA 94070 (US). **WHYTE, David** [US/US]; 2623 Barclay Way, Belmont, CA 94002 (US). **MANNING, Gerard** [IE/US]; 844 Fremont Street #4, Menlo Park, CA 94025 (US). **SUDARSANAM, Sucha** [US/US]; 20 Corte Patencio, Greenbrae, CA 94904 (US). **MARTINEZ, Ricardo** [US/US]; 984 Cartier Lane, Foster City, CA 94404 (US).

(74) Agents: **BURROUS, Beth, A.** et al.; Foley & Lardner, 3000 K Street N.W., Suite 500, Washington, DC 20007-5109 (US).

(81) Designated States *(national)*: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/66594 A2

(54) Title: NOVEL HUMAN PROTEIN KINASES AND PROTEIN KINASE-LIKE ENZYMES

(57) Abstract: The present invention relates to kinase polypeptides, nucleotide sequences encoding the kinase polypeptides, as well as various products and methods useful for the diagnosis and treatment of various kinase-related diseases and conditions. Through the use of a bioinformatics strategy, mammalian members of the PTK's and STK's have been identified and their protein structure predicted.

NOVEL HUMAN PROTEIN KINASES AND PROTEIN KINASE-LIKE ENZYMES

The present invention claims priority on provisional application serial nos. 60/187,150, and 60/193,404, and 60/247,103 all of which are hereby incorporated by reference in their entirety.

FIELD OF THE INVENTION

The present invention relates to kinase polypeptides, nucleotide sequences encoding the kinase polypeptides, as well as various products and methods useful for the diagnosis and treatment of various kinase-related diseases and conditions.

BACKGROUND OF THE INVENTION

The following description of the background of the invention is provided to aid in understanding the invention, but is not admitted to be or to describe prior art to the invention.

Cellular signal transduction is a fundamental mechanism whereby external stimuli that regulate diverse cellular processes are relayed to the interior of cells. One of the key biochemical mechanisms of signal transduction involves the reversible phosphorylation of proteins, which enables regulation of the activity of mature proteins by altering their structure and function.

Protein phosphorylation plays a pivotal role in cellular signal transduction. Among the biological functions controlled by this type of postranslational modification are: cell division, differentiation and death (apoptosis); cell motility and cytoskeletal structure; control of DNA replication, transcription, splicing and translation; protein translocation events from the endoplasmic reticulum and Golgi apparatus to the membrane and extracellular space; protein nuclear import and export; regulation of metabolic reactions, etc. Abnormal protein phosphorylation is widely recognized to be causally linked to the etiology of many diseases including cancer as well as immunologic, neuronal and metabolic disorders.

The following abbreviations are used for kinases throught this application:

ASK Apoptosis signal-regulating kinase

CaMK Ca2+/calmodulin-dependent protein kinase

CCRK	Cell cycle-related kinase
CDK	Cyclin-dependent kinase
CK	Casein kinase
DAPK	Death-associated protein kinase
DM	myotonic dystrophy kinase
Dyrk	dual-specificity-tyrosine phosphorylating-regulated kinase
GAK	Cyclin G-associated kinase
GRK	G-protein coupled receptor
GuC	Guanylate cyclase
HIPK	Homeodomain-interacting protein kinase
IRAK	Interleukin-1 receptor-associated kinase
MAPK	Mitogen activated protein kinase
MAST	Microtubule-associated STK
MLCK	Myosin-light chain kinase
MLK	Mixed lineage kinase
NIMA	NimA-related protein kinase
PKA	cAMP-dependent protein kinase
RSK	Ribosomal protein S6 kinase
RTK	Receptor tyrosine kinase
SGK	Serum and glucocorticoid-regulated kinase
STK	serine threonine kinase
ULK	UNC-51-like kinase

The best-characterized protein kinases in eukaryotes phosphorylate proteins on the hydroxyl substituent of serine, threonine and tyrosine residues, which are the most common phospho-acceptor amino acid residues. However, phosphorylation on histidine has also been observed in bacteria.

The presence of a phosphate moiety modulates protein function in multiple ways. A common mechanism includes changes in the catalytic properties (Vmax and Km) of an enzyme, leading to its activation or inactivation.

A second widely recognized mechanism involves promoting protein-protein interactions. An example of this is the tyrosine autophosphorylation of the ligand-activated EGF receptor tyrosine kinase. This event triggers the high-affinity binding to the phosphotyrosine residue on the receptor's C-terminal intracellular domain to

the SH2 motif of the adaptor molecule Grb2. Grb2, in turn, binds through its SH3 motif to a second adaptor molecule, such as SHC. The formation of this ternary complex activates the signaling events that are responsible for the biological effects of EGF. Serine and threonine phosphorylation events also have been recently recognized to exert their biological function through protein-protein interaction events that are mediated by the high-affinity binding of phosphoserine and phosphothreonine to WW motifs present in a large variety of proteins (Lu, P.J. et al (1999) Science 283:1325-1328).

A third important outcome of protein phosphorylation is changes in the subcellular localization of the substrate. As an example, nuclear import and export events in a large diversity of proteins are regulated by protein phosphorylation (Drier E.A. et al (1999) Genes Dev 13: 556-568).

Protein kinases are one of the largest families of eukaryotic proteins with several hundred known members. These proteins share a 250-300 amino acid domain that can be subdivided into 12 distinct subdomains that comprise the common catalytic core structure. These conserved protein motifs have recently been exploited using PCR-based and bioinformatic strategies leading to a significant expansion of the known kinases. Multiple alignment of the sequences in the catalytic domain of protein kinases and subsequent parsimony analysis permits their segregation into sub-families of related kinases.

Kinases largely fall into two groups: those specific for phosphorylating serines and threonines, and those specific for phosphorylating tyrosines. Some kinases, referred to as "dual specificity" kinases, are able to phosphorylate on tyrosine as well as serine/threonine residues.

Protein kinases can also be characterized by their location within the cell. Some kinases are transmembrane receptor-type proteins capable of directly altering their catalytic activity in response to the external environment such as the binding of a ligand. Others are non-receptor-type proteins lacking any transmembrane domain. They can be found in a variety of cellular compartments from the inner surface of the cell membrane to the nucleus.

Many kinases are involved in regulatory cascades wherein their substrates may include other kinases whose activities are regulated by their phosphorylation state. Ultimately the activity of some downstream effector is modulated by phosphorylation resulting from activation of such a pathway. The conserved protein motifs of these

kinases have recently been exploited using PCR-based cloning strategies leading to a significant expansion of the known kinases.

Multiple alignment of the sequences in the catalytic domain of protein kinases and subsequent parsimony analysis permits the segregation of related kinases into distinct branches of subfamilies including: tyrosine kinases (PTK's), dual-specificity kinases, and serine/threonine kinases (STK's). The latter subfamily includes cyclic-nucleotide-dependent kinases, calcium/calmodulin kinases, cyclin-dependent kinases (CDK's), MAP-kinases, serine-threonine kinase receptors, and several other less defined subfamilies.

The protein kinases may be classified into several major groups including AGC, CAMK, Casein kinase 1, CMGC, STE, tyrosine kinases, and atypical kinases (Plowman, GD *et al.*, *Proceedings of the National Academy of Sciences*, USA, Vol. 96, Issue 24, 13603-13610, November 23, 1999; see also www.kinase.com). In addition, there are a number of minor yet distinct families, including families related to worm- or fungal-specific kinases, and a family designated "other" to represent several smaller families. Within each group are several distinct families of more closely related kinases. In addition, an "atypical" family represents those protein kinases whose catalytic domain has little or no primary sequence homology to conventional kinases, including the A6 kinases and PI3 kinases.

AGC group

The AGC kinases are basic amino acid-directed enzymes that phosphorylate residues found proximal to Arg and Lys. Examples of this group are the G protein-coupled receptor kinases (GRKs), the cyclic nucleotide-dependent kinases (PKA, PKC, PKG), NDR or DBF2 kinases, ribosomal S6 kinases, AKT kinases, myotonic dystrophy kinases (DMPKs), MAPK interacting kinases (MNKs), MAST kinases, and Mo3C11.1_ce family originally identified only in nematodes.

GRKs regulate signaling from heterotrimeric guanine protein coupled receptors (GPCRs). Mutations in GPCRs cause a number of human diseases, including retinitis pigmentosa, stationary night blindness, color blindness , hyperfunctioning thyroid adenomas, familial precocious puberty , familial hypocalciuric hypercalcemia and neonatal severe hyperparathroidism (OMIM, http://www.ncbi.nlm.nih.gov/Omim/). The regulation of GPCRs by GRKs indirectly implicates GRKs in these diseases.

The cAMP-dependent protein kinases (PKA) consist of heterotetramers comprised of 2 catalytic (C) and 2 regulatory (R) subunits, in which the R subunits bind to the second messenger cAMP, leading to dissociation of the active C subunits from the complex. Many of these kinases respond to second messengers such as cAMP resulting in a wide range of cellular responses to hormones and neurotransmitters.

AKT is a mammalian proto-oncoprotein regulated by phosphatidylinositol 3-kinase (PI3-K), which appears to function as a cell survival signal to protect cells from apoptosis. Insulin receptor, RAS, PI3-K, and PDK1 all act as upstream activators of AKT, whereas the lipid phosphatase PTEN functions as a negative regulator of the PI3-K/AKT pathway. Downstream targets for AKT-mediated cell survival include the pro-apoptotic factors BAD and Caspase9, and transcription factors in the forkhead family, such as DAF-16 in the worm. AKT is also an essential mediator in insulin signaling, in part due to its use of GSK-3 as another downstream target.

The S6 kinases regulate a wide array of cellular processes involved in mitogenic response including protein synthesis, translation of specific mRNA species, and cell cycle progression from G1 to S phase. The gene has been localized to chromosomal region 17q23 and is amplified in breast cancer (Couch, *et al.*, Cancer Res. 1999 Apr 1;59(7):1408-11).

CAMK Group

The CAMK kinases are also basic amino acid-directed kinases. They include the Ca2+/calmodulin-regulated and AMP-dependent protein kinases (AMPK), myosin light chain kinases (MLCK), MAP kinase activating protein kinases (MAPKAPKs) checkpoint 2 kinases (CHK2), death-associated protein kinases (DAPKs), phosphorylase kinase (PHK), Rac and Rho-binding Trio kinases, a "unique" family of CAMKs, and the EMK-related protein kinases.

The EMK family of STKs are involved in the control of cell polarity, microtubule stability and cancer. One member of the EMK family, C-TAK1, has been reported to control entry into mitosis by activating Cdc25C which in turn dephosphorylates Cdc2. Also included in the EMK family is MAKV, which has been shown to be overexpressed in metastatic tumors (*Dokl. Akad. Nauk* 354 (4), 554-556 (1997)).

Tyrosine Protein Kinase Group (TK)

The tyrosine kinase group encompass both cytoplasmic (*e.g.* src) as well as transmembrane receptor tyrosine kinases (*e.g.* EGF receptor). These kinases play a pivotal role in the signal transduction processes that mediate cell proliferation, differentiation and apoptosis.

Casein Kinase 1 Group (CK1)

The CK1 family represents a distant branch of the protein kinase family. The hallmarks of protein kinase subdomains VIII and IX are difficult to identify. One or more forms are ubiquitously distributed in mammalian tissues and cell lines. CK1 kinases are found in cytoplasm, in nuclei, membrane-bound, and associated with the cytoskeleton. Splice variants differ in their subcellular distribution.

"Other" Group

Several families cluster within a group of unrelated kinases termed "Other". Included are: CHK1; Elongation 2 factor kinases (EIFK); homologues of the yeast sterile family kinases (STE), which refers to 3 classes of kinases which lie sequentially upstream of the MAPKs; Calcium-calmodulin kinase kinases (CAMKK); dual-specific tyrosine kinases (DYRK); IkB kinases (IKK); Integrin receptor kinase (IRAK); endoribonuclease-associated kinases (IRE); Mixed lineage kinase (MLK); LIM-domain containing kinase (LIMK); MOS; PIM; Receptor interacting kinase (RIP); SR-protein specific kinase (SRPK); RAF; Serine-threonine kinase receptors (STKR); TAK1; Testis specific kinase (TSK); tousled-related kinase (TSL); UNC51-related kinase (UNC); VRK; WEE; mitotic kinases (BUB1, AURORA, PLK, and NIMA/NEK); several families that are close homologues to worm (C26C2.1, YQ09, ZC581.9, YFL033c, C24A1.3); Drosophila (SLOB), or yeast (YDOD_sp, YGR262_sc) kinases; and others that are "unique," that is, those which do not cluster into any obvious family. Additional families are even less well defined and first were identified in lower eukaryotes such as yeast or worms (YNL020, YPL236, YQ09, YWY3, SCY1, C01H6.9, C26C2.1)

RIP2 is a serine-threonine kinase associated with the tumor necrosis factor (TNF) receptor complex and is implicated in the activation of NF-kappa B and cell death in mammalian cells. It has recently been demonstrated that RIP2 activates the

MAPK pathway (Navas, *et al.*, *J Biol. Chem.* 1999 Nov 19;274(47):33684-33690). RIP2 activates AP-1 and serum response element regulated expression by inducing the activation of the Elk1 transcription factor. RIP2 directly phosphorylates and activates ERK2 *in vivo* and *in vitro*. RIP2 in turn is activated through its interaction with Ras-activated Raf1. These results highlight the integrated nature of kinase signaling pathway.

The tousled (TSL) kinase was first identified in the plant Arabidopsis thaliana. TSL encodes a serine/threonine kinase that is essential for proper flower development. Human tousled-like kinases (Tlks) are cell-cycle-regulated enzymes, displaying maximal activities during S phase. This regulated activity suggests that Tlk function is linked to ongoing DNA replication (Sillje, *et al.*, *EMBO J* 1999 Oct 15;18(20):5691-5702).

Atypical Protein Kinase Group

There are several proteins with protein kinase activity that appear structurally unrelated to the eukaryotic protein kinases. These include; *Dictyostelium* myosin heavy chain kinase A (MHCKA), *Physarum polycephalum* actin-fragmin kinase, the human A6 PTK, human BCR, mitochondrial pyruvate dehydrogenase and branched chain fatty acid dehydrogenase kinase, and the prokaryotic "histidine" protein kinase family. The slime mold, worm, and human eEF-2 kinase homologues have all been demonstrated to have protein kinase activity, yet they bear little resemblance to conventional protein kinases except for the presence of a putative GxGxxG ATP-binding motif.

The so-called histidine kinases are abundant in prokaryotes, with more than 20 representatives in *E. coli*, and have also been identified in yeast, molds, and plants. In response to external stimuli, these kinases act as part of two-component systems to regulate DNA replication, cell division, and differentiation through phosphorylation of an aspartate in the target protein. To date, no "histidine" kinases have been identified in metazoans, although mitochondrial pyruvate dehydrogenase (PDK) and branched chain alpha-ketoacid dehydrogenase kinase (BCKD kinase), are related in sequence. PDK and BCKD kinase represent a unique family of atypical protein kinases involved in regulation of glycolysis, the citric acid cycle, and protein synthesis during protein malnutrition. Structurally they conserve only the C-terminal portion of "histidine" kinases including the G box regions. BCKD kinase

phosphorylates the E1a subunit of the BCKD complex on Ser-293, proving it to be a functional protein kinase. Although no bona fide "histidine" kinase has yet been identified in humans, they do contain PDK.

Several other proteins contain protein kinase-like homology including: receptor guanylyl cyclases, diacylglycerol kinases, choline/ethanolamine kinases, and YLK1-related antibiotic resistance kinases. Each of these families contain short motifs that were recognized by our profile searches with low scoring E-values, but a *priori* would not be expected to function as protein kinases. Instead, the similarity could simply reflect the modular nature of protein evolution and the primal role of ATP binding in diverse phosphotransfer enzymes. However, two recent papers on a bacterial homologue of the YLK1 family suggests that the aminoglycoside phosphotransferases (APHs) are structurally and functionally related to protein kinases. There are over 40 APHs identified from bacteria that are resistant to aminoglycosides such as kanamycin, gentamycin, or amikacin. The crystal structure of one well characterized APH reveals that it shares greater than 40% structural identity with the 2 lobed structure of the catalytic domain of cAMP-dependent protein kinase (PKA), including an N-terminal lobe composed of a 5-stranded antiparallel beta sheet and the core of the C-terminal lobe including several invariant segments found in all protein kinases. APHs lack the GxGxxG normally present in the loop between beta strands 1 and 2 but contain 7 of the 12 strictly conserved residues present in most protein kinases, including the HGDxxxN signature sequence in kinase subdomain VIB. Furthermore, APH also has been shown to exhibit protein-serine/threonine kinase activity, suggesting that other YLK-related molecules may indeed be functional protein kinases.

The eukaryotic lipid kinases (PI3Ks, PI4Ks, and PIPKs) also contain several short motifs similar to protein kinases, but otherwise share minimal primary sequence similarity. However, once again structural analysis of PIPKII-beta defines a conserved ATP-binding core that is strikingly similar to conventional protein kinases. Three residues are conserved among all of these enzymes including (relative to the PKA sequence) Lys-72 which binds the gamma-phosphate of ATP, Asp-166 which is part of the HRDLK motif and Asp-184 from the conserved Mg^{++} or Mn^{++} binding DFG motif. The worm genome contains 12 phosphatidylinositol kinases, including 3 PI3-kinases, 2 PI4-kinases, 3 PIP5-kinases, and 4 PI3-kinase-related kinases. The latter group has 4 mammalian members (DNA-PK, FRAP/TOR, ATM, and ATR), which

have been shown to participate in the maintenance of genomic integrity in response to DNA damage, and exhibit true protein kinase activity, raising the possibility that other PI-kinases may also act as protein kinases. Regardless of whether they have true protein kinase activity, PI3-kinases are tightly linked to protein kinase signaling, as evidenced by their involvement downstream of many growth factor receptors and as upstream activators of the cell survival response mediated by the AKT protein kinase.

SUMMARY OF THE INVENTION

The present invention relates, in part, to human protein kinases and protein kinase-like enzymes identified from genomic sequencing.

Tyrosine and serine/threonine kinases (PTK's and STK's) have been identified and their protein sequence predicted as part of the instant invention. Mammalian members of these families were identified through the use of a bioinformatics strategy. The partial or complete sequences of these kinases are presented here, together with their classification, predicted or deduced protein structure.

One aspect of the invention features an identified, isolated, enriched, or purified nucleic acid molecule encoding a kinase polypeptide and having a nucleic acid sequence selected from the group consisting of those set forth in SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, and SEQ ID NO: 12.

The term "identified" in reference to a nucleic acid is meant that a sequence was selected from a genomic, EST, or cDNA sequence database based on it being predicted to encode a portion of a previously unknown or novel protein kinase.

By "isolated," in reference to nucleic acid, is meant a polymer of 10 (preferably 21, more preferably 39, most preferably 75) or more nucleotides conjugated to each other, including DNA and RNA that is isolated from a natural source or that is synthesized as the sense or complementary antisense strand. In certain embodiments of the invention, longer nucleic acids are preferred, for example those of 300, 600, 900, 1200, 1500, or more nucleotides and/or those having at least 50%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95% or 99% identity to a sequence selected from the group consisting of those set forth in SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID

NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, and SEQ ID NO: 12.

The isolated nucleic acid of the present invention is unique in the sense that it is not found in a pure or separated state in nature. Use of the term "isolated" indicates that a naturally occurring sequence has been removed from its normal cellular (*i.e.*, chromosomal) environment. Thus, the sequence may be in a cell-free solution or placed in a different cellular environment. The term does not imply that the sequence is the only nucleotide chain present, but that it is essentially free (about 90 - 95% pure at least) of non-nucleotide material naturally associated with it, and thus is distinguished from isolated chromosomes.

By the use of the term "enriched" in reference to nucleic acid is meant that the specific DNA or RNA sequence constitutes a significantly higher fraction (2- to 5-fold) of the total DNA or RNA present in the cells or solution of interest than in normal or diseased cells or in the cells from which the sequence was taken. This could be caused by a person by preferential reduction in the amount of other DNA or RNA present, or by a preferential increase in the amount of the specific DNA or RNA sequence, or by a combination of the two. However, it should be noted that enriched does not imply that there are no other DNA or RNA sequences present, just that the relative amount of the sequence of interest has been significantly increased. The term "significant" is used to indicate that the level of increase is useful to the person making such an increase, and generally means an increase relative to other nucleic acids of about at least 2-fold, more preferably at least 5- to 10-fold or even more. The term also does not imply that there is no DNA or RNA from other sources. The DNA from other sources may, for example, comprise DNA from a yeast or bacterial genome, or a cloning vector such as pUC19. This term distinguishes from naturally occurring events, such as viral infection, or tumor-type growths, in which the level of one mRNA may be naturally increased relative to other species of mRNA. That is, the term is meant to cover only those situations in which a person has intervened to elevate the proportion of the desired nucleic acid.

It is also advantageous for some purposes that a nucleotide sequence be in purified form. The term "purified" in reference to nucleic acid does not require absolute purity (such as a homogeneous preparation). Instead, it represents an indication that the sequence is relatively more pure than in the natural environment (compared to the natural level this level should be at least 2- to 5-fold greater, *e.g.*, in

terms of mg/mL). Individual clones isolated from a cDNA library may be purified to electrophoretic homogeneity. The claimed DNA molecules obtained from these clones could be obtained directly from total DNA or from total RNA. The cDNA clones are not naturally occurring, but rather are preferably obtained via manipulation of a partially purified naturally occurring substance (messenger RNA). The construction of a cDNA library from mRNA involves the creation of a synthetic substance (cDNA) and pure individual cDNA clones can be isolated from the synthetic library by clonal selection of the cells carrying the cDNA library. Thus, the process which includes the construction of a cDNA library from mRNA and isolation of distinct cDNA clones yields an approximately 10^6-fold purification of the native message. Thus, purification of at least one order of magnitude, preferably two or three orders, and more preferably four or five orders of magnitude is expressly contemplated.

By a "kinase polypeptide" is meant 32 (preferably 40, more preferably 45, most preferably 55) or more contiguous amino acids in a polypeptide having an amino acid sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24. In certain aspects, polypeptides of 100, 200, 300, 400, 450, 500, 550, 600, 700, 800, 900 or more amino acids are preferred. The kinase polypeptide can be encoded by a full-length nucleic acid sequence or any portion (*e.g.*, a "fragment" as defined herein) of the full-length nucleic acid sequence, so long as a functional activity of the polypeptide is retained, including, for example, a catalytic domain, as defined herein, or a portion thereof. One of skill in the art would be able to select those catalytic domains, or portions thereof, which exhibit a kinase or kinase-like activity, *e.g.*, catalytic activity, as defined herein. It is well known in the art that due to the degeneracy of the genetic code numerous different nucleic acid sequences can code for the same amino acid sequence. Equally, it is also well known in the art that conservative changes in amino acid can be made to arrive at a protein or polypeptide which retains the functionality of the original. Such substitutions may include the replacement of an amino acid by a residue having similar physicochemical properties, such as substituting one aliphatic residue (Ile, Val, Leu or Ala) for another, or substitution between basic residues Lys and Arg, acidic residues Glu and Asp, amide residues Gln and Asn, hydroxyl residues Ser and Tyr, or aromatic residues Phe

and Tyr. Further information regarding making amino acid exchanges which have only slight, if any, effects on the overall protein can be found in Bowie *et al.*, *Science*, 1990, *247*, 1306-1310, which is incorporated herein by reference in its entirety including any figures, tables, or drawings. In all cases, all permutations are intended to be covered by this disclosure.

The amino acid sequence of a kinase peptide of the invention will be substantially similar to a sequence having an amino acid sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24, or the corresponding full-length amino acid sequence, or fragments thereof.

A sequence that is substantially similar to a sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24, will preferably have at least 90% identity (more preferably at least 95% and most preferably 99-100%) to the sequence.

By "identity" is meant a property of sequences that measures their similarity or relationship. Identity is measured by dividing the number of identical residues by the total number of residues and gaps and multiplying the product by 100. "Gaps" are spaces in an alignment that are the result of additions or deletions of amino acids. Thus, two copies of exactly the same sequence have 100% identity, but sequences that are less highly conserved, and have deletions, additions, or replacements, may have a lower degree of identity. Those skilled in the art will recognize that several computer programs are available for determining sequence identity using standard parameters, for example Gapped BLAST or PSI-BLAST (Altschul, *et al.* (1997) *Nucleic Acids Res.* 25:3389-3402), BLAST (Altschul, *et al.* (1990) *J. Mol. Biol.* 215:403-410), and Smith-Waterman (Smith, *et al.* (1981) *J. Mol. Biol.* 147:195-197). Preferably, the default settings of these programs will be employed, but those skilled in the art recognize whether these settings need to be changed and know how to make the changes.

"Similarity" is measured by dividing the number of identical residues plus the number of conservatively substituted residues (see Bowie, *et al. Science*, 1999), 247, 1306-1310, which is incorporated herein by reference in its entirety, including any

drawings, figures, or tables) by the total number of residues and gaps and multiplying the product by 100.

In preferred embodiments, the invention features isolated, enriched, or purified nucleic acid molecules encoding a kinase polypeptide comprising a nucleotide sequence that: (a) encodes a polypeptide having an amino acid sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24; (b) is the complement of the nucleotide sequence of (a); (c) hybridizes under highly stringent conditions to the nucleotide molecule of (a) and encodes a naturally occurring kinase polypeptide; (d) encodes a polypeptide having an amino acid sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24, except that it lacks one or more, but not all, of the domains selected from the group consisting of an N-terminal domain, a catalytic domain, a C-terminal catalytic domain, a C-terminal domain, a coiled-coil structure region, a proline-rich region, a spacer region, and a C-terminal tail; and (e) is the complement of the nucleotide sequence of (d).

The term "complement" refers to two nucleotides that can form multiple favorable interactions with one another. For example, adenine is complementary to thymine as they can form two hydrogen bonds. Similarly, guanine and cytosine are complementary since they can form three hydrogen bonds. A nucleotide sequence is the complement of another nucleotide sequence if all of the nucleotides of the first sequence are complementary to all of the nucleotides of the second sequence.

Various low or high stringency hybridization conditions may be used depending upon the specificity and selectivity desired. These conditions are well known to those skilled in the art. Under stringent hybridization conditions only highly complementary nucleic acid sequences hybridize. Preferably, such conditions prevent hybridization of nucleic acids having more than 1 or 2 mismatches out of 20 contiguous nucleotides, more preferably, such conditions prevent hybridization of nucleic acids having more than 1 or 2 mismatches out of 50 contiguous nucleotides, most preferably, such conditions prevent hybridization of nucleic acids having more than 1 or 2 mismatches out of 100 contiguous nucleotides. In some instances, the

conditions may prevent hybridization of nucleic acids having more than 5 mismatches in the full-length sequence.

By stringent hybridization assay conditions is meant hybridization assay conditions at least as stringent as the following: hybridization in 50% formamide, 5X SSC, 50 mM NaH2PO4, pH 6.8, 0.5% SDS, 0.1 mg/mL sonicated salmon sperm DNA, and 5X Denhardt's solution at 42 °C overnight; washing with 2X SSC, 0.1% SDS at 45 °C; and washing with 0.2X SSC, 0.1% SDS at 45 °C. Under some of the most stringent hybridization assay conditions, the second wash can be done with 0.1X SSC at a temperature up to 70 °C (Berger *et al.* (1987) Guide to Molecular Cloning Techniques pg 421, hereby incorporated by reference herein in its entirety including any figures, tables, or drawings.). However, other applications may require the use of conditions falling between these sets of conditions. Methods of determining the conditions required to achieve desired hybridizations are well known to those with ordinary skill in the art, and are based on several factors, including but not limited to, the sequences to be hybridized and the samples to be tested. Washing conditions of lower stringency frequently utilize a lower temperature during the washing steps, such as 65 °C, 60 °C, 55 °C, 50 °C, or 42 °C.

The term "domain" refers to a region of a polypeptide which serves a particular function. For instance, N-terminal or C-terminal domains of signal transduction proteins can serve functions including, but not limited to, binding molecules that localize the signal transduction molecule to different regions of the cell or binding other signaling molecules directly responsible for propagating a particular cellular signal. Some domains can be expressed separately from the rest of the protein and function by themselves, while others must remain part of the intact protein to retain function. The latter are termed functional regions of proteins and also relate to domains.

The term "N-terminal domain" refers to the extracatalytic region located between the initiator methionine and the catalytic domain of the protein kinase. The N-terminal domain can be identified following a Smith-Waterman alignment of the protein sequence against the non-redundant protein database to define the N-terminal boundary of the catalytic domain. Depending on its length, the N-terminal domain may or may not play a regulatory role in kinase function. An example of a protein kinase whose N-terminal domain has been shown to play a regulatory role is PAK65,

which contains a CRIB motif used for Cdc42 and rac binding (Burbelo, P.D. *et al.* (1995) *J. Biol. Chem.* 270, 29071-29074).

The term "catalytic domain" refers to a region of the protein kinase that is typically 25-300 amino acids long and is responsible for carrying out the phosphate transfer reaction from a high-energy phosphate donor molecule such as ATP or GTP to itself (autophosphorylation) or to other proteins (exogenous phosphorylation). The catalytic domain of protein kinases is made up of 12 subdomains that contain highly conserved amino acid residues, and are responsible for proper polypeptide folding and for catalysis. The catalytic domain can be identified following a Smith-Waterman alignment of the protein sequence against the non-redundant protein database.

The term "catalytic activity", as used herein, defines the rate at which a kinase catalytic domain phosphorylates a substrate. Catalytic activity can be measured, for example, by determining the amount of a substrate converted to a phosphorylated product as a function of time. Catalytic activity can be measured by methods of the invention by holding time constant and determining the concentration of a phosphorylated substrate after a fixed period of time. Phosphorylation of a substrate occurs at the active site of a protein kinase. The active site is normally a cavity in which the substrate binds to the protein kinase and is phosphorylated.

The term "substrate" as used herein refers to a molecule phosphorylated by a kinase of the invention. Kinases phosphorylate substrates on serine/threonine or tyrosine amino acids. The molecule may be another protein or a polypeptide.

The term "C-terminal domain" refers to the region located between the catalytic domain or the last (located closest to the C-terminus) functional domain and the carboxy-terminal amino acid residue of the protein kinase. By "functional" domain is meant any region of the polypeptide that may play a regulatory or catalytic role as predicted from amino acid sequence homology to other proteins or by the presence of amino acid sequences that may give rise to specific structural conformations (*e.g.* N-terminal domain). The C-terminal domain can be identified by using a Smith-Waterman alignment of the protein sequence against the non-redundant protein database to define the C-terminal boundary of the catalytic domain or of any functional C-terminal extracatalytic domain. Depending on its length and amino acid composition, the C-terminal domain may or may not play a regulatory role in kinase function. An example of a protein kinase whose C-terminal domain may play a regulatory role is PAK3 which contains a heterotrimeric G_b subunit-binding site near

its C-terminus (Leeuw, T. *et al.* (1998) *Nature*, 391, 191-195). For the some of the kinases of the instant invention, the C-terminal domain may also comprise the catalytic domain (above).

The term "C-terminal tail" as used herein, refers to a C-terminal domain of a protein kinase, that by homology extends or protrudes past the C-terminal amino acid of its closest homolog. C-terminal tails can be identified by using a Smith-Waterman sequence alignment of the protein sequence against the non-redundant protein database, or by means of a multiple sequence alignment of homologous sequences using the DNAStar program Megalign. Depending on its length, a C-terminal tail may or may not play a regulatory role in kinase function.

The term "coiled-coil structure region" as used herein, refers to a polypeptide sequence that has a high probability of adopting a coiled-coil structure as predicted by computer algorithms such as COILS (Lupas, A. (1996) *Meth. Enzymology* 266:513-525). Coiled-coils are formed by two or three amphipathic α-helices in parallel. Coiled-coils can bind to coiled-coil domains of other polypeptides resulting in homo- or heterodimers (Lupas, A. (1991) *Science* 252:1162-1164). Coiled-coil-dependent oligomerization has been shown to be necessary for protein function including catalytic activity of serine/threonine kinases (Roe, J. *et al.* (1997) *J. Biol. Chem.* 272:5838-5845).

The term "proline-rich region" as used herein, refers to a region of a protein kinase whose proline content over a given amino acid length is higher than the average content of this amino acid found in proteins(*i.e.*, >10%). Proline-rich regions are easily discernable by visual inspection of amino acid sequences and quantitated by standard computer sequence analysis programs such as the DNAStar program EditSeq. Proline-rich regions have been demonstrated to participate in regulatory protein -protein interactions. Among these interactions, those that are most relevant to this invention involve the "PxxP" proline rich motif found in certain protein kinases (*i.e.*, human PAK1) and the SH3 domain of the adaptor molecule Nck (Galisteo, M.L. *et al.* (1996) *J. Biol. Chem.* 271:20997-21000). Other regulatory interactions involving "PxxP" proline-rich motifs include the WW domain (Sudol, M. (1996) *Prog. Biochys. Mol. Bio.* 65:113-132).

The term "spacer region" as used herein, refers to a region of the protein kinase located between predicted functional domains. The spacer region has no

detectable homology to any amino acid sequence in the database, and can be identified by using a Smith-Waterman alignment of the protein sequence against the non-redundant protein database to define the C- and N-terminal boundaries of the flanking functional domains. Spacer regions may or may not play a fundamental role in protein kinase function. Precedence for the regulatory role of spacer regions in kinase function is provided by the role of the *src* kinase spacer in inter-domain interactions (Xu, W. *et al.* (1997) *Nature* 385:595-602).

The term "insert" as used herein refers to a portion of a protein kinase that is absent from a close homolog. Inserts may or may not by the product alternative splicing of exons. Inserts can be identified by using a Smith-Waterman sequence alignment of the protein sequence against the non-redundant protein database, or by means of a multiple sequence alignment of homologous sequences using the DNAStar program Megalign. Inserts may play a functional role by presenting a new interface for protein-protein interactions, or by interfering with such interactions.

The term "signal transduction pathway" refers to the molecules that propagate an extracellular signal through the cell membrane to become an intracellular signal. This signal can then stimulate a cellular response. The polypeptide molecules involved in signal transduction processes are typically receptor and non-receptor protein tyrosine kinases, receptor and non-receptor protein phosphatases, polypeptides containing *SRC* homology 2 and 3 domains, phosphotyrosine binding proteins (*SRC* homology 2 (SH2) and phosphotyrosine binding (PTB and PH) domain containing proteins), proline-rich binding proteins (SH3 domain containing proteins), GTPases, phosphodiesterases, phospholipases, prolyl isomerases, proteases, Ca2+ binding proteins, cAMP binding proteins, guanyl cyclases, adenylyl cyclases, NO generating proteins, nucleotide exchange factors, and transcription factors.

In other preferred embodiments, the invention features isolated, enriched, or purified nucleic acid molecules encoding kinase polypeptides, further comprising a vector or promoter effective to initiate transcription in a host cell. The invention also features recombinant nucleic acid, preferably in a cell or an organism. The recombinant nucleic acid may contain a sequence selected from the group consisting of those set forth in SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, and SEQ ID NO: 12, or a functional derivative thereof and a vector or a promoter effective to initiate transcription in a host cell. The

recombinant nucleic acid can alternatively contain a transcriptional initiation region functional in a cell, a sequence complementary to an RNA sequence encoding a kinase polypeptide and a transcriptional termination region functional in a cell. Specific vectors and host cell combinations are discussed herein.

The term "vector" relates to a single or double-stranded circular nucleic acid molecule that can be transfected into cells and replicated within or independently of a cell genome. A circular double-stranded nucleic acid molecule can be cut and thereby linearized upon treatment with restriction enzymes. An assortment of nucleic acid vectors, restriction enzymes, and the knowledge of the nucleotide sequences cut by restriction enzymes are readily available to those skilled in the art. A nucleic acid molecule encoding a kinase can be inserted into a vector by cutting the vector with restriction enzymes and ligating the two pieces together.

The term "transfecting" defines a number of methods to insert a nucleic acid vector or other nucleic acid molecules into a cellular organism. These methods involve a variety of techniques, such as treating the cells with high concentrations of salt, an electric field, detergent, or DMSO to render the outer membrane or wall of the cells permeable to nucleic acid molecules of interest or use of various viral transduction strategies.

The term "promoter" as used herein, refers to nucleic acid sequence needed for gene sequence expression. Promoter regions vary from organism to organism, but are well known to persons skilled in the art for different organisms. For example, in prokaryotes, the promoter region contains both the promoter (which directs the initiation of RNA transcription) as well as the DNA sequences which, when transcribed into RNA, will signal synthesis initiation. Such regions will normally include those 5'-non-coding sequences involved with initiation of transcription and translation, such as the TATA box, capping sequence, CAAT sequence, and the like.

In preferred embodiments, the isolated nucleic acid comprises, consists essentially of, or consists of a nucleic acid sequence selected from the group consisting of those set forth in SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, and SEQ ID NO: 12, which encodes an amino acid sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID

NO: 23, and SEQ ID NO: 24, a functional derivative thereof, or at least 35, 40, 45, 50, 60, 75, 100, 200, or 300 contiguous amino acids selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24. The nucleic acid may be isolated from a natural source by cDNA cloning or by subtractive hybridization. The natural source may be mammalian, preferably human, preferably blood, semen or tissue, and the nucleic acid may be synthesized by the triester method or by using an automated DNA synthesizer.

The term "mammal" refers preferably to such organisms as mice, rats, rabbits, guinea pigs, sheep, and goats, more preferably to cats, dogs, monkeys, and apes, and most preferably to humans.

In yet other preferred embodiments, the nucleic acid is a conserved or unique region, for example those useful for: the design of hybridization probes to facilitate identification and cloning of additional polypeptides, the design of PCR probes to facilitate cloning of additional polypeptides, obtaining antibodies to polypeptide regions, and designing antisense oligonucleotides.

By "conserved nucleic acid regions", are meant regions present on two or more nucleic acids encoding a kinase polypeptide, to which a particular nucleic acid sequence can hybridize under lower stringency conditions. Examples of lower stringency conditions suitable for screening for nucleic acid encoding kinase polypeptides are provided in Wahl *et al. Meth. Enzym.* 152:399-407 (1987) and in Wahl *et al. Meth. Enzym.* 152:415-423 (1987), which are hereby incorporated by reference herein in its entirety, including any drawings, figures, or tables. Preferably, conserved regions differ by no more than 5 out of 20 nucleotides, even more preferably 2 out of 20 nucleotides or most preferably 1 out of 20 nucleotides.

By "unique nucleic acid region" is meant a sequence present in a nucleic acid coding for a kinase polypeptide that is not present in a sequence coding for any other naturally occurring polypeptide. Such regions preferably encode 32 (preferably 40, more preferably 45, most preferably 55) or more contiguous amino acids, for example, an amino acid sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID

NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24. In particular, a unique nucleic acid region is preferably of mammalian origin.

Another aspect of the invention features a nucleic acid probe for the detection of nucleic acid encoding a kinase polypeptide having an amino acid sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24 in a sample. The nucleic acid probe contains a nucleotide base sequence that will hybridize to the sequence selected from the group consisting of those set forth in SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, and SEQ ID NO: 12, or a functional derivative thereof.

In preferred embodiments, the nucleic acid probe hybridizes to nucleic acid encoding at least 12, 32, 75, 90, 105, 120, 150, 200, 250, 300 or 350 contiguous amino acids, wherein the nucleic acid sequence is selected from the group consisting of SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, and SEQ ID NO: 12, or a functional derivative thereof.

Methods for using the probes include detecting the presence or amount of kinase RNA in a sample by contacting the sample with a nucleic acid probe under conditions such that hybridization occurs and detecting the presence or amount of the probe bound to kinase RNA. The nucleic acid duplex formed between the probe and a nucleic acid sequence coding for a kinase polypeptide may be used in the identification of the sequence of the nucleic acid detected (Nelson *et al.*, *in Nonisotopic DNA Probe Techniques*, Academic Press, San Diego, Kricka, ed., p. 275, 1992, hereby incorporated by reference herein in its entirety, including any drawings, figures, or tables). Kits for performing such methods may be constructed to include a container means having disposed therein a nucleic acid probe.

Methods for using the probes also include using these probes to find, for example, the full-length clone of each of the predicted kinases by techniques known to one skilled in the art. These clones will be useful for screening for small molecule compounds that inhibit the catalytic activity of the encoded kinase with potential utility in treating cancers, immune-related diseases and disorders, cardiovascular disease, brain or neuronal-associated diseases, and metabolic disorders. More

specifically disorders including cancers of tissues, blood, or hematopoietic origin, particularly those involving breast, colon, lung, prostate, cervical, brain, ovarian, bladder, or kidney; central or peripheral nervous system diseases and conditions including migraine, pain, sexual dysfunction, mood disorders, attention disorders, cognition disorders, hypotension, and hypertension; psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias, such as Huntington's disease or Tourette's Syndrome; neurodegenerative diseases including Alzheimer's, Parkinson's, multiple sclerosis, and amyotrophic lateral sclerosis; viral or non-viral infections caused by HIV-1, HIV-2 or other viral- or prion-agents or fungal- or bacterial- organisms; metabolic disorders including Diabetes and obesity and their related syndromes, among others; cardiovascular disorders including reperfusion restenosis, coronary thrombosis, clotting disorders, unregulated cell growth disorders, atherosclerosis; ocular disease including glaucoma, retinopathy, and macular degeneration; inflammatory disorders including rheumatoid arthritis, chronic inflammatory bowel disease, chronic inflammatory pelvic disease, multiple sclerosis, asthma, osteoarthritis, psoriasis, atherosclerosis, rhinitis, autoimmunity, and organ transplant rejection.

In another aspect, the invention describes a recombinant cell or tissue comprising a nucleic acid molecule encoding a kinase polypeptide having an amino acid sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24. In such cells, the nucleic acid may be under the control of the genomic regulatory elements, or may be under the control of exogenous regulatory elements including an exogenous promoter. By "exogenous" it is meant a promoter that is not normally coupled *in vivo* transcriptionally to the coding sequence for the kinase polypeptides.

The polypeptide is preferably a fragment of the protein encoded by an amino acid sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24. By "fragment," is meant an amino acid sequence present in a kinase polypeptide. Preferably, such a sequence comprises at least 32, 45, 50, 60,

100, 200, or 300 contiguous amino acids of a sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24.

In another aspect, the invention features an isolated, enriched, or purified kinase polypeptide having the amino acid sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24.

By "isolated" in reference to a polypeptide is meant a polymer of 6 (preferably 12, more preferably 18, most preferably 25, 32, 40, or 50) or more amino acids conjugated to each other, including polypeptides that are isolated from a natural source or that are synthesized. In certain aspects longer polypeptides are preferred, such as those comprising 100, 200, 300, 400, 450, 500, 550, 600, 700, 800, 900 or more contiguous amino acids, including an amino acid sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24.

The isolated polypeptides of the present invention are unique in the sense that they are not found in a pure or separated state in nature. Use of the term "isolated" indicates that a naturally occurring sequence has been removed from its normal cellular environment. Thus, the sequence may be in a cell-free solution or placed in a different cellular environment. The term does not imply that the sequence is the only amino acid chain present, but that it is essentially free (about 90 - 95% pure at least) of non-amino acid-based material naturally associated with it.

By the use of the term "enriched" in reference to a polypeptide is meant that the specific amino acid sequence constitutes a significantly higher fraction (2- to 5-fold) of the total amino acid sequences present in the cells or solution of interest than in normal or diseased cells or in the cells from which the sequence was taken. This could be caused by a person by preferential reduction in the amount of other amino acid sequences present, or by a preferential increase in the amount of the specific amino acid sequence of interest, or by a combination of the two. However, it should be noted that enriched does not imply that there are no other amino acid sequences present, just that the relative amount of the sequence of interest has been significantly

increased. The term "significantly" here is used to indicate that the level of increase is useful to the person making such an increase, and generally means an increase relative to other amino acid sequences of about at least 2-fold, more preferably at least 5- to 10-fold or even more. The term also does not imply that there is no amino acid sequence from other sources. The other source of amino acid sequences may, for example, comprise amino acid sequence encoded by a yeast or bacterial genome, or a cloning vector such as pUC19. The term is meant to cover only those situations in which man has intervened to increase the proportion of the desired amino acid sequence.

It is also advantageous for some purposes that an amino acid sequence be in purified form. The term "purified" in reference to a polypeptide does not require absolute purity (such as a homogeneous preparation); instead, it represents an indication that the sequence is relatively purer than in the natural environment. Compared to the natural level this level should be at least 2-to 5-fold greater (*e.g.*, in terms of mg/mL). Purification of at least one order of magnitude, preferably two or three orders, and more preferably four or five orders of magnitude is expressly contemplated. The substance is preferably free of contamination at a functionally significant level, for example 90%, 95%, or 99% pure.

In preferred embodiments, the kinase polypeptide is a fragment of the protein encoded by an amino acid sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24. Preferably, the kinase polypeptide contains at least 32, 45, 50, 60, 100, 200, or 300 contiguous amino acids of a sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24, or a functional derivative thereof.

In preferred embodiments, the kinase polypeptide comprises an amino acid sequence having (a) an amino acid sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24, except that it lacks one or more of the domains selected from the group consisting of a C-terminal catalytic

domain, an N-terminal domain, a catalytic domain, a C-terminal domain, a coiled-coil structure region, a proline-rich region, a spacer region, and a C-terminal tail.

The polypeptide can be isolated from a natural source by methods well-known in the art. The natural source may be mammalian, preferably human, preferably blood, semen or tissue, and the polypeptide may be synthesized using an automated polypeptide synthesizer.

In some embodiments the invention includes a recombinant kinase polypeptide having (a) an amino acid sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24. By "recombinant kinase polypeptide" is meant a polypeptide produced by recombinant DNA techniques such that it is distinct from a naturally occurring polypeptide either in its location (*e.g.*, present in a different cell or tissue than found in nature), purity or structure. Generally, such a recombinant polypeptide will be present in a cell in an amount different from that normally observed in nature.

The polypeptides to be expressed in host cells may also be fusion proteins which include regions from heterologous proteins. Such regions may be included to allow, *e.g.*, secretion, improved stability, or facilitated purification of the polypeptide. For example, a sequence encoding an appropriate signal peptide can be incorporated into expression vectors. A DNA sequence for a signal peptide (secretory leader) may be fused in-frame to the polynucleotide sequence so that the polypeptide is translated as a fusion protein comprising the signal peptide. A signal peptide that is functional in the intended host cell promotes extracellular secretion of the polypeptide. Preferably, the signal sequence will be cleaved from the polypeptide upon secretion of the polypeptide from the cell. Thus, preferred fusion proteins can be produced in which the N-terminus of a kinase polypeptide is fused to a carrier peptide.

In one embodiment, the polypeptide comprises a fusion protein which includes a heterologous region used to facilitate purification of the polypeptide. Many of the available peptides used for such a function allow selective binding of the fusion protein to a binding partner. A preferred binding partner includes one or more of the IgG binding domains of protein A are easily purified to homogeneity by affinity chromatography on, for example, IgG-coupled Sepharose. Alternatively, many vectors have the advantage of carrying a stretch of histidine residues that can be

expressed at the N-terminal or C-terminal end of the target protein, and thus the protein of interest can be recovered by metal chelation chromatography. A nucleotide sequence encoding a recognition site for a proteolytic enzyme such as enterokinase, factor X procollagenase or thrombine may immediately precede the sequence for a kinase polypeptide to permit cleavage of the fusion protein to obtain the mature kinase polypeptide. Additional examples of fusion-protein binding partners include, but are not limited to, the yeast I-factor, the honeybee melatin leader in sf9 insect cells, 6-His tag, thioredoxin tag, hemaglutinin tag, GST tag, and OmpA signal sequence tag. As will be understood by one of skill in the art, the binding partner which recognizes and binds to the peptide may be any ion, molecule or compound including metal ions (*e.g.*, metal affinity columns), antibodies, or fragments thereof, and any protein or peptide which binds the peptide, such as the FLAG tag.

In another aspect, the invention features an antibody (*e.g.*, a monoclonal or polyclonal antibody) having specific binding affinity to a kinase polypeptide or a kinase polypeptide domain or fragment where the polypeptide is selected from the group having an amino acid sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24. By "specific binding affinity" is meant that the antibody binds to the target kinase polypeptide with greater affinity than it binds to other polypeptides under specified conditions. Antibodies or antibody fragments are polypeptides that contain regions that can bind other polypeptides. The term "specific binding affinity" describes an antibody that binds to a kinase polypeptide with greater affinity than it binds to other polypeptides under specified conditions.. Antibodies can be used to identify an endogenous source of kinase polypeptides, to monitor cell cycle regulation, and for immuno-localization of kinase polypeptides within the cell.

The term "polyclonal" refers to antibodies that are heterogenous populations of antibody molecules derived from the sera of animals immunized with an antigen or an antigenic functional derivative thereof. For the production of polyclonal antibodies, various host animals may be immunized by injection with the antigen. Various adjuvants may be used to increase the immunological response, depending on the host species.

"Monoclonal antibodies" are substantially homogenous populations of antibodies to a particular antigen. They may be obtained by any technique which provides for the production of antibody molecules by continuous cell lines in culture. Monoclonal antibodies may be obtained by methods known to those skilled in the art (Kohler *et al.*, *Nature* 256:495-497, 1975, and U.S. Patent No. 4,37 6,110, both of which are hereby incorporated by reference herein in their entirety including any figures, tables, or drawings).

The term "antibody fragment" refers to a portion of an antibody, often the hypervariable region and portions of the surrounding heavy and light chains, that displays specific binding affinity for a particular molecule. A hypervariable region is a portion of an antibody that physically binds to the polypeptide target.

Antibodies or antibody fragments having specific binding affinity to a kinase polypeptide of the invention may be used in methods for detecting the presence and/or amount of kinase polypeptide in a sample by probing the sample with the antibody under conditions suitable for kinase-antibody immunocomplex formation and detecting the presence and/or amount of the antibody conjugated to the kinase polypeptide. Diagnostic kits for performing such methods may be constructed to include antibodies or antibody fragments specific for the kinase as well as a conjugate of a binding partner of the antibodies or the antibodies themselves.

An antibody or antibody fragment with specific binding affinity to a kinase polypeptide of the invention can be isolated, enriched, or purified from a prokaryotic or eukaryotic organism. Routine methods known to those skilled in the art enable production of antibodies or antibody fragments, in both prokaryotic and eukaryotic organisms. Purification, enrichment, and isolation of antibodies, which are polypeptide molecules, are described above.

Antibodies having specific binding affinity to a kinase polypeptide of the invention may be used in methods for detecting the presence and/or amount of kinase polypeptide in a sample by contacting the sample with the antibody under conditions such that an immunocomplex forms and detecting the presence and/or amount of the antibody conjugated to the kinase polypeptide. Diagnostic kits for performing such methods may be constructed to include a first container containing the antibody and a second container having a conjugate of a binding partner of the antibody and a label, such as, for example, a radioisotope. The diagnostic kit may also include notification of an FDA approved use and instructions therefor.

In another aspect, the invention features a hybridoma which produces an antibody having specific binding affinity to a kinase polypeptide or a kinase polypeptide domain, where the polypeptide is selected from the group having an amino acid sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24. By "hybridoma" is meant an immortalized cell line that is capable of secreting an antibody, for example an antibody to a kinase of the invention. In preferred embodiments, the antibody to the kinase comprises a sequence of amino acids that is able to specifically bind a kinase polypeptide of the invention.

In another aspect, the present invention is also directed to kits comprising antibodies that bind to a polypeptide encoded by any of the nucleic acid molecules described above, and a negative control antibody.

The term "negative control antibody" refers to an antibody derived from similar source as the antibody having specific binding affinity, but where it displays no binding affinity to a polypeptide of the invention.

In another aspect, the invention features a kinase polypeptide binding agent able to bind to a kinase polypeptide selected from the group having (a) an amino acid sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24. The binding agent is preferably a purified antibody that recognizes an epitope present on a kinase polypeptide of the invention. Other binding agents include molecules that bind to kinase polypeptides and analogous molecules that bind to a kinase polypeptide. Such binding agents may be identified by using assays that measure kinase binding partner activity, such as those that measure PDGFR activity.

The invention also features a method for screening for human cells containing a kinase polypeptide of the invention or an equivalent sequence. The method involves identifying the novel polypeptide in human cells using techniques that are routine and standard in the art, such as those described herein for identifying the kinases of the invention (*e.g.*, cloning, Southern or Northern blot analysis, *in situ* hybridization, PCR amplification, etc.).

In another aspect, the invention features methods for identifying a substance that modulates kinase activity comprising the steps of: (a) contacting a kinase polypeptide selected from the group having an amino acid sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24 with a test substance; (b) measuring the activity of said polypeptide; and (c) determining whether said substance modulates the activity of said polypeptide. The skilled artisan will appreciate that the kinase polypeptides of the invention, including, for example, a portion of a full-length sequence such as a catalytic domain or a portion thereof, are useful for the identification of a substance which modulates kinase activity. Those kinase polypeptides having a functional activity (*e.g.*, catalytic activity as defined herein) are useful for identifying a substance that modulates kinase activity.

The term "modulates" refers to the ability of a compound to alter the function of a kinase of the invention. A modulator preferably activates or inhibits the activity of a kinase of the invention depending on the concentration of the compound exposed to the kinase.

The term "modulates" also refers to altering the function of kinases of the invention by increasing or decreasing the probability that a complex forms between the kinase and a natural binding partner. A modulator preferably increases the probability that such a complex forms between the kinase and the natural binding partner, more preferably increases or decreases the probability that a complex forms between the kinase and the natural binding partner depending on the concentration of the compound exposed to the kinase, and most preferably decreases the probability that a complex forms between the kinase and the natural binding partner.

The term "activates" refers to increasing the cellular activity of the kinase. The term inhibit refers to decreasing the cellular activity of the kinase. Kinase activity is preferably the interaction with a natural binding partner.

The term "complex" refers to an assembly of at least two molecules bound to one another. Signal transduction complexes often contain at least two protein molecules bound to one another. For instance, a protein tyrosine receptor protein kinase, GRB2, SOS, RAF, and RAS assemble to form a signal transduction complex in response to a mitogenic ligand.

The term "natural binding partner" refers to polypeptides, lipids, small molecules, or nucleic acids that bind to kinases in cells. A change in the interaction between a kinase and a natural binding partner can manifest itself as an increased or decreased probability that the interaction forms, or an increased or decreased concentration of kinase/natural binding partner complex.

The term "contacting" as used herein refers to mixing a solution comprising the test compound with a liquid medium bathing the cells of the methods. The solution comprising the compound may also comprise another component, such as dimethyl sulfoxide (DMSO), which facilitates the uptake of the test compound or compounds into the cells of the methods. The solution comprising the test compound may be added to the medium bathing the cells by utilizing a delivery apparatus, such as a pipette-based device or syringe-based device.

In another aspect, the invention features methods for identifying a substance that modulates kinase activity in a cell comprising the steps of: (a) expressing a kinase polypeptide in a cell, wherein said polypeptide is selected from the group having an amino acid sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24; (b) adding a test substance to said cell; and (c) monitoring a change in cell phenotype or the interaction between said polypeptide and a natural binding partner. The skilled artisan will appreciate that the kinase polypeptides of the invention, including, for example, a portion of a full-length sequence such as a catalytic domain or a portion thereof, are useful for the identification of a substance which modulates kinase activity. Those kinase polypeptides having a functional activity (*e.g.*, catalytic activity as defined herein) are useful for identifying a substance that modulates kinase activity.

The term "expressing" as used herein refers to the production of kinases of the invention from a nucleic acid vector containing kinase genes within a cell. The nucleic acid vector is transfected into cells using well known techniques in the art as described herein.

Another aspect of the instant invention is directed to methods of identifying compounds that bind to kinase polypeptides of the present invention, comprising contacting the kinase polypeptides with a compound, and determining whether the compound binds the kinase polypeptides. Binding can be determined by binding

assays which are well known to the skilled artisan, including, but not limited to, gel-shift assays, Western blots, radiolabeled competition assay, phage-based expression cloning, co-fractionation by chromatography, co-precipitation, cross linking, interaction trap/two-hybrid analysis, southwestern analysis, ELISA, and the like, which are described in, for example, *Current Protocols in Molecular Biology*, 1999, John Wiley & Sons, NY, which is incorporated herein by reference in its entirety. The compounds to be screened include, but are not limited to, compounds of extracellular, intracellular, biological or chemical origin.

The methods of the invention also embrace compounds that are attached to a label, such as a radiolabel (*e.g.*, ^{125}I, ^{35}S, ^{32}P, ^{33}P, ^{3}H), a fluorescence label, a chemiluminescent label, an enzymic label and an immunogenic label. The kinase polypeptides employed in such a test may either be free in solution, attached to a solid support, borne on a cell surface, located intracellularly or associated with a portion of a cell. One skilled in the art can, for example, measure the formation of complexes between a kinase polypeptide and the compound being tested. Alternatively, one skilled in the art can examine the diminution in complex formation between a kinase polypeptide and its substrate caused by the compound being tested.

Other assays can be used to examine enzymatic activity including, but not limited to, photometric, radiometric, HPLC, electrochemical, and the like, which are described in, for example, *Enzyme Assays: A Practical Approach*, eds. R. Eisenthal and M. J. Danson, 1992, Oxford University Press, which is incorporated herein by reference in its entirety.

Another aspect of the present invention is directed to methods of identifying compounds which modulate (*i.e.*, increase or decrease) activity of a kinase polypeptide comprising contacting the kinase polypeptide with a compound, and determining whether the compound modifies activity of the kinase polypeptide. As described herein, the kinase polypeptides of the invention include a portion of a full-length sequence, such as a catalytic domain, as defined herein. In some instances, the kinase polypeptides of the invention comprise less than the entire catalytic domain, yet exhibit kinase or kinase-like activity. These compounds are also referred to as "modulators of protein kinases." The activity in the presence of the test compound is measured to the activity in the absence of the test compound. Where the activity of a sample containing the test compound is higher than the activity in a sample lacking the test compound, the compound will have increased the activity. Similarly, where

the activity of a sample containing the test compound is lower than the activity in the sample lacking the test compound, the compound will have inhibited the activity.

The present invention is particularly useful for screening compounds by using a kinase polypeptide in any of a variety of drug screening techniques. The compounds to be screened include, but are not limited to, extracellular, intracellular, biological or chemical origin. The kinase polypeptide employed in such a test may be in any form, preferably, free in solution, attached to a solid support, borne on a cell surface or located intracellularly. One skilled in the art can, for example, measure the formation of complexes between a kinase polypeptide and the compound being tested. Alternatively, one skilled in the art can examine the diminution in complex formation between a kinase polypeptide and its substrate caused by the compound being tested.

The activity of kinase polypeptides of the invention can be determined by, for example, examining the ability to bind or be activated by chemically synthesised peptide ligands. Alternatively, the activity of the kinase polypeptides can be assayed by examining their ability to bind metal ions such as calcium, hormones, chemokines, neuropeptides, neurotransmitters, nucleotides, lipids, odorants, and photons. Thus, modulators of the kinase polypeptide's activity may alter a kinase function, such as a binding property of a kinase or an activity such as signal transduction or membrane localization.

In various embodiments of the method, the assay may take the form of a yeast growth assay, an Aequorin assay, a Luciferase assay, a mitogenesis assay, a MAP Kinase activity assay, as well as other binding or function-based assays of kinase activity that are generally known in the art. In several of these embodiments, the invention includes any of the receptor and non-receptor protein tyrosine kinases, receptor and non-receptor protein phosphatases, polypeptides containing *SRC* homology 2 and 3 domains, phosphotyrosine binding proteins (*SRC* homology 2 (SH2) and phosphotyrosine binding (PTB and PH) domain containing proteins), proline-rich binding proteins (SH3 domain containing proteins), GTPases, phosphodiesterases, phospholipases, prolyl isomerases, proteases, Ca2+ binding proteins, cAMP binding proteins, guanyl cyclases, adenylyl cyclases, NO generating proteins, nucleotide exchange factors, and transcription factors. Biological activities of kinases according to the invention include, but are not limited to, the binding of a natural or a synthetic ligand, as well as any one of the functional activities of kinases known in the art. Non-limiting examples of kinase activities include transmembrane

signaling of various forms, which may involve kinase binding interactions and/or the exertion of an influence over signal transduction.

The modulators of the invention exhibit a variety of chemical structures, which can be generally grouped into mimetics of natural kinase ligands, and peptide and non-peptide allosteric effectors of kinases. The invention does not restrict the sources for suitable modulators, which may be obtained from natural sources such as plant, animal or mineral extracts, or non-natural sources such as small molecule libraries, including the products of combinatorial chemical approaches to library construction, and peptide libraries.

The use of cDNAs encoding kinases in drug discovery programs is well-known; assays capable of testing thousands of unknown compounds per day in high-throughput screens (HTSs) are thoroughly documented. The literature is replete with examples of the use of radiolabelled ligands in HTS binding assays for drug discovery (see Williams, *Medicinal Research Reviews*, 1991, *11*, 147-184.; Sweetnam, *et al.*, *J. Natural Products*, 1993, *56*, 441-455 for review). Recombinant receptors are preferred for binding assay HTS because they allow for better specificity (higher relative purity), provide the ability to generate large amounts of receptor material, and can be used in a broad variety of formats (see Hodgson, *Bio/Technology*, 1992, *10*, 973-980; each of which is incorporated herein by reference in its entirety).

A variety of heterologous systems is available for functional expression of recombinant receptors that are well known to those skilled in the art. Such systems include bacteria (Strosberg, *et al.*, *Trends in Pharmacological Sciences*, 1992, *13*, 95-98), yeast (Pausch, *Trends in Biotechnology*, 1997, *15*, 487-494), several kinds of insect cells (Vanden Broeck, *Int. Rev. Cytology*, 1996, *164*, 189-268), amphibian cells (Jayawickreme *et al.*, *Current Opinion in Biotechnology*, 1997, *8*, 629-634) and several mammalian cell lines (CHO, HEK293, COS, etc.; see Gerhardt, *et al.*, *Eur. J. Pharmacology*, 1997, *334*, 1-23). These examples do not preclude the use of other possible cell expression systems, including cell lines obtained from nematodes (PCT application WO 98/37177).

An expressed kinase can be used for HTS binding assays in conjunction with its defined ligand, in this case the corresponding peptide that activates it. The identified peptide is labeled with a suitable radioisotope, including, but not limited to, ^{125}I, ^{3}H, ^{35}S or ^{32}P, by methods that are well known to those skilled in the art. Alternatively, the peptides may be labeled by well-known methods with a suitable

fluorescent derivative (Baindur, *et al., Drug Dev. Res.*, 1994, 33, 373-398; Rogers, *Drug Discovery Today*, 1997, *2*, 156-160). Radioactive ligand specifically bound to the receptor in membrane preparations made from the cell line expressing the recombinant protein can be detected in HTS assays in one of several standard ways, including filtration of the receptor-ligand complex to separate bound ligand from unbound ligand (Williams, *Med. Res. Rev.*, 1991, *11*, 147-184.; Sweetnam, *et al., J. Natural Products*, 1993, 56, 441-455). Alternative methods include a scintillation proximity assay (SPA) or a FlashPlate format in which such separation is unnecessary (Nakayama, *Cur. Opinion Drug Disc. Dev.*, 1998, *1*, 85-91 Bossé, *et al., J. Biomolecular Screening*, 1998, *3*, 285-292.). Binding of fluorescent ligands can be detected in various ways, including fluorescence energy transfer (FRET), direct spectrophotofluorometric analysis of bound ligand, or fluorescence polarization (Rogers, *Drug Discovery Today*, 1997, *2*, 156-160; Hill, *Cur. Opinion Drug Disc. Dev.*, 1998, 1, 92-97).

The kinases and natural binding partners required for functional expression of heterologous kinase polypeptides can be native constituents of the host cell or can be introduced through well-known recombinant technology. The kinase polypeptides can be intact or chimeric. The kinase activation results in the stimulation or inhibition of other native proteins, events that can be linked to a measurable response.

Examples of such biological responses include, but are not limited to, the following: the ability to survive in the absence of a limiting nutrient in specifically engineered yeast cells (Pausch, *Trends in Biotechnology*, 1997, *15*, 487-494); changes in intracellular Ca^{2+} concentration as measured by fluorescent dyes (Murphy, *et al., Cur. Opinion Drug Disc. Dev.*, 1998, *1*, 192-199). Fluorescence changes can also be used to monitor ligand-induced changes in membrane potential or intracellular pH; an automated system suitable for HTS has been described for these purposes (Schroeder, *et al., J. Biomolecular Screening*, 1996, *1*, 75-80). Assays are also available for the measurement of common second but these are not generally preferred for HTS.

The invention contemplates a multitude of assays to screen and identify inhibitors of ligand binding to kinase polypeptides. In one example, the kinase polypeptide is immobilized and interaction with a binding partner is assessed in the presence and absence of a candidate modulator such as an inhibitor compound. In another example, interaction between the kinase polypeptide and its binding partner is assessed in a solution assay, both in the presence and absence of a candidate inhibitor

compound. In either assay, an inhibitor is identified as a compound that decreases binding between the kinase polypeptide and its natural binding partner. Another contemplated assay involves a variation of the di-hybrid assay wherein an inhibitor of protein/protein interactions is identified by detection of a positive signal in a transformed or transfected host cell, as described in PCT publication number WO 95/20652, published August 3, 1995 and is included by reference herein including any figures, tables, or drawings.

Candidate modulators contemplated by the invention include compounds selected from libraries of either potential activators or potential inhibitors. There are a number of different libraries used for the identification of small molecule modulators, including: (1) chemical libraries, (2) natural product libraries, and (3) combinatorial libraries comprised of random peptides, oligonucleotides or organic molecules. Chemical libraries consist of random chemical structures, some of which are analogs of known compounds or analogs of compounds that have been identified as "hits" or "leads" in other drug discovery screens, while others are derived from natural products, and still others arise from non-directed synthetic organic chemistry. Natural product libraries are collections of microorganisms, animals, plants, or marine organisms which are used to create mixtures for screening by: (1) fermentation and extraction of broths from soil, plant or marine microorganisms or (2) extraction of plants or marine organisms. Natural product libraries include polyketides, non-ribosomal peptides, and variants (non-naturally occurring) thereof. For a review, see *Science* 282:63-68 (1998). Combinatorial libraries are composed of large numbers of peptides, oligonucleotides, or organic compounds as a mixture. These libraries are relatively easy to prepare by traditional automated synthesis methods, PCR, cloning, or proprietary synthetic methods. Of particular interest are non-peptide combinatorial libraries. Still other libraries of interest include peptide, protein, peptidomimetic, multiparallel synthetic collection, recombinatorial, and polypeptide libraries. For a review of combinatorial chemistry and libraries created therefrom, see Myers, *Curr. Opin. Biotechnol.* 8:701-707 (1997). Identification of modulators through use of the various libraries described herein permits modification of the candidate "hit" (or "lead") to optimize the capacity of the "hit" to modulate activity.

Still other candidate inhibitors contemplated by the invention can be designed and include soluble forms of binding partners, as well as such binding partners as chimeric, or fusion, proteins. A "binding partner" as used herein broadly

encompasses both natural binding partners as described above as well as chimeric polypeptides, peptide modulators other than natural ligands, antibodies, antibody fragments, and modified compounds comprising antibody domains that are immunospecific for the expression product of the identified kinase gene.

Other assays may be used to identify specific peptide ligands of a kinase polypeptide, including assays that identify ligands of the target protein through measuring direct binding of test ligands to the target protein, as well as assays that identify ligands of target proteins through affinity ultrafiltration with ion spray mass spectroscopy/HPLC methods or other physical and analytical methods. Alternatively, such binding interactions are evaluated indirectly using the yeast two-hybrid system described in Fields *et al.*, *Nature*, 340:245-246 (1989), and Fields *et al.*, *Trends in Genetics*, 10:286-292 (1994), both of which are incorporated herein by reference. The two-hybrid system is a genetic assay for detecting interactions between two proteins or polypeptides. It can be used to identify proteins that bind to a known protein of interest, or to delineate domains or residues critical for an interaction. Variations on this methodology have been developed to clone genes that encode DNA binding proteins, to identify peptides that bind to a protein, and to screen for drugs. The two-hybrid system exploits the ability of a pair of interacting proteins to bring a transcription activation domain into close proximity with a DNA binding domain that binds to an upstream activation sequence (UAS) of a reporter gene, and is generally performed in yeast. The assay requires the construction of two hybrid genes encoding (1) a DNA-binding domain that is fused to a first protein and (2) an activation domain fused to a second protein. The DNA-binding domain targets the first hybrid protein to the UAS of the reporter gene; however, because most proteins lack an activation domain, this DNA-binding hybrid protein does not activate transcription of the reporter gene. The second hybrid protein, which contains the activation domain, cannot by itself activate expression of the reporter gene because it does not bind the UAS. However, when both hybrid proteins are present, the noncovalent interaction of the first and second proteins tethers the activation domain to the UAS, activating transcription of the reporter gene. For example, when the first protein is a kinase gene product, or fragment thereof, that is known to interact with another protein or nucleic acid, this assay can be used to detect agents that interfere with the binding interaction. Expression of the reporter gene is monitored as different test agents are added to the system. The presence of an inhibitory agent results in lack of a reporter signal.

When the function of the kinase polypeptide gene product is unknown and no ligands are known to bind the gene product, the yeast two-hybrid assay can also be used to identify proteins that bind to the gene product. In an assay to identify proteins that bind to a kinase polypeptide, or fragment thereof, a fusion polynucleotide encoding both a kinase polypeptide (or fragment) and a UAS binding domain (*i.e.*, a first protein) may be used. In addition, a large number of hybrid genes each encoding a different second protein fused to an activation domain are produced and screened in the assay. Typically, the second protein is encoded by one or more members of a total cDNA or genomic DNA fusion library, with each second protein coding region being fused to the activation domain. This system is applicable to a wide variety of proteins, and it is not even necessary to know the identity or function of the second binding protein. The system is highly sensitive and can detect interactions not revealed by other methods; even transient interactions may trigger transcription to produce a stable mRNA that can be repeatedly translated to yield the reporter protein.

Other assays may be used to search for agents that bind to the target protein. One such screening method to identify direct binding of test ligands to a target protein is described in U.S. Patent No. 5,585,277, incorporated herein by reference. This method relies on the principle that proteins generally exist as a mixture of folded and unfolded states, and continually alternate between the two states. When a test ligand binds to the folded form of a target protein (*i.e.*, when the test ligand is a ligand of the target protein), the target protein molecule bound by the ligand remains in its folded state. Thus, the folded target protein is present to a greater extent in the presence of a test ligand which binds the target protein, than in the absence of a ligand. Binding of the ligand to the target protein can be determined by any method which distinguishes between the folded and unfolded states of the target protein. The function of the target protein need not be known in order for this assay to be performed. Virtually any agent can be assessed by this method as a test ligand, including, but not limited to, metals, polypeptides, proteins, lipids, polysaccharides, polynucleotides and small organic molecules.

Another method for identifying ligands of a target protein is described in Wieboldt *et al.*, *Anal. Chem.*, 69:1683-1691 (1997), incorporated herein by reference. This technique screens combinatorial libraries of 20-30 agents at a time in solution phase for binding to the target protein. Agents that bind to the target protein are separated from other library components by simple membrane washing. The

specifically selected molecules that are retained on the filter are subsequently liberated from the target protein and analyzed by HPLC and pneumatically assisted electrospray (ion spray) ionization mass spectroscopy. This procedure selects library components with the greatest affinity for the target protein, and is particularly useful for small molecule libraries.

In preferred embodiments of the invention, methods of screening for compounds which modulate kinase activity comprise contacting test compounds with kinase polypeptides and assaying for the presence of a complex between the compound and the kinase polypeptide. In such assays, the ligand is typically labelled. After suitable incubation, free ligand is separated from that present in bound form, and the amount of free or uncomplexed label is a measure of the ability of the particular compound to bind to the kinase polypeptide.

In another embodiment of the invention, high throughput screening for compounds having suitable binding affinity to kinase polypeptides is employed. Briefly, large numbers of different small peptide test compounds are synthesised on a solid substrate. The peptide test compounds are contacted with the kinase polypeptide and washed. Bound kinase polypeptide is then detected by methods well known in the art. Purified polypeptides of the invention can also be coated directly onto plates for use in the aforementioned drug screening techniques. In addition, non-neutralizing antibodies can be used to capture the protein and immobilize it on the solid support.

Other embodiments of the invention comprise using competitive screening assays in which neutralizing antibodies capable of binding a polypeptide of the invention specifically compete with a test compound for binding to the polypeptide. In this manner, the antibodies can be used to detect the presence of any peptide that shares one or more antigenic determinants with a kinase polypeptide. Radiolabeled competitive binding studies are described in A.H. Lin *et al. Antimicrobial Agents and Chemotherapy*, 1997, vol. 41, no. 10. pp. 2127-2131, the disclosure of which is incorporated herein by reference in its entirety.

In another aspect, the invention provides methods for treating a disease by administering to a patient in need of such treatment a substance that modulates the activity of a kinase polypeptide selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24, as well as the full-length polypeptide thereof,

or a portion of any of these sequences that retains functional activity, as described herein. Preferably the disease is selected from the group consisting of cancers, immune-elated diseases and disorders, cardiovascular disease, brain or neuronal-associated diseases, and metabolic disorders. More specifically these diseases include cancer of tissues, blood, or hematopoietic origin, particularly those involving breast, colon, lung, prostate, cervical, brain, ovarian, bladder, or kidney; central or peripheral nervous system diseases and conditions including migraine, pain, sexual dysfunction, mood disorders, attention disorders, cognition disorders, hypotension, and hypertension; psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias, such as Huntington's disease or Tourette's Syndrome; neurodegenerative diseases including Alzheimer's, Parkinson's, Multiple sclerosis, and Amyotrophic lateral sclerosis; viral or non-viral infections caused by HIV-1, HIV-2 or other viral- or prion-agents or fungal- or bacterial- organisms; metabolic disorders including Diabetes and obesity and their related syndromes, among others; cardiovascular disorders including reperfusion restenosis, coronary thrombosis, clotting disorders, unregulated cell growth disorders, atherosclerosis; ocular disease including glaucoma, retinopathy, and macular degeneration; inflammatory disorders including rheumatoid arthritis, chronic inflammatory bowel disease, chronic inflammatory pelvic disease, multiple sclerosis, asthma, osteoarthritis, psoriasis, atherosclerosis, rhinitis, autoimmunity, and organ transplant rejection.

In preferred embodiments, the invention provides methods for treating or preventing a disease or disorder by administering to a patient in need of such treatment a substance that modulates the activity of a kinase polypeptide having an amino acid sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24, as well as the full-length polypeptide thereof, or a portion of any of these sequences that retains functional activity, as described herein. Preferably, the disease is selected from the group consisting of cancers, immune-related diseases and disorders, cardiovascular disease, brain or neuronal-associated diseases, and metabolic disorders. More specifically these diseases include cancer of tissues, blood, or hematopoietic origin, particularly those involving breast, colon, lung, prostate, cervical, brain, ovarian, bladder, or kidney; central or peripheral

nervous system diseases and conditions including migraine, pain, sexual dysfunction, mood disorders, attention disorders, cognition disorders, hypotension, and hypertension; psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias, such as Huntington's disease or Tourette's Syndrome; neurodegenerative diseases including Alzheimer's, Parkinson's, Multiple sclerosis, and Amyotrophic lateral sclerosis; viral or non-viral infections caused by HIV-1, HIV-2 or other viral- or prion-agents or fungal- or bacterial- organisms; metabolic disorders including Diabetes and obesity and their related syndromes, among others; cardiovascular disorders including reperfusion restenosis, coronary thrombosis, clotting disorders, unregulated cell growth disorders, atherosclerosis; ocular disease including glaucoma, retinopathy, and macular degeneration; inflammatory disorders including rheumatoid arthritis, chronic inflammatory bowel disease, chronic inflammatory pelvic disease, multiple sclerosis, asthma, osteoarthritis, psoriasis, atherosclerosis, rhinitis, autoimmunity, and organ transplant rejection.

The invention also features methods of treating or preventing a disease or disorder by administering to a patient in need of such treatment a substance that modulates the activity of a kinase polypeptide having an amino acid sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24, as well as the full-length polypeptide thereof, or a portion of any of these sequences that retains functional activity, as described herein. Preferably the disease is selected from the group consisting of cancers, immune-related diseases and disorders, cardiovascular disease, brain or neuronal-associated diseases, and metabolic disorders. More specifically these diseases include cancer of tissues, blood, or hematopoietic origin, particularly those involving breast, colon, lung, prostate, cervical, brain, ovarian, bladder, or kidney; central or peripheral nervous system diseases and conditions including migraine, pain, sexual dysfunction, mood disorders, attention disorders, cognition disorders, hypotension, and hypertension; psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias, such as Huntington's disease or Tourette's Syndrome; neurodegenerative diseases including Alzheimer's, Parkinson's, Multiple sclerosis, and Amyotrophic lateral sclerosis; viral or non-viral infections

caused by HIV-1, HIV-2 or other viral- or prion-agents or fungal- or bacterial-organisms; metabolic disorders including Diabetes and obesity and their related syndromes, among others; cardiovascular disorders including reperfusion restenosis, coronary thrombosis, clotting disorders, unregulated cell growth disorders, atherosclerosis; ocular disease including glaucoma, retinopathy, and macular degeneration; inflammatory disorders including rheumatoid arthritis, chronic inflammatory bowel disease, chronic inflammatory pelvic disease, multiple sclerosis, asthma, osteoarthritis, psoriasis, atherosclerosis, rhinitis, autoimmunity, and organ transplant rejection.

The invention also features methods of treating or preventing a disease or disorder by administering to a patient in need of such treatment a substance that modulates the activity of a kinase polypeptide having an amino acid sequence selected from the group consisting those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24, as well as the full-length polypeptide thereof, or a portion of any of these sequences that retains functional activity, as described herein. Preferably the disease is selected from the group consisting of immune-related diseases and disorders, cardiovascular disease, and cancer. More preferably these diseases include cancer of tissues, blood, or hematopoietic origin, particularly those involving breast, colon, lung, prostate, cervical, brain, ovarian, bladder, or kidney; central or peripheral nervous system diseases and conditions including migraine, pain, sexual dysfunction, mood disorders, attention disorders, cognition disorders, hypotension, and hypertension; psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias, such as Huntington's disease or Tourette's Syndrome; neurodegenerative diseases including Alzheimer's, Parkinson's, Multiple sclerosis, and Amyotrophic lateral sclerosis; viral or non-viral infections caused by HIV-1, HIV-2 or other viral- or prion-agents or fungal- or bacterial-organisms; metabolic disorders including Diabetes and obesity and their related syndromes, among others; cardiovascular disorders including reperfusion restenosis, coronary thrombosis, clotting disorders, unregulated cell growth disorders, atherosclerosis; ocular disease including glaucoma, retinopathy, and macular degeneration; inflammatory disorders including rheumatoid arthritis, chronic inflammatory bowel disease, chronic inflammatory pelvic disease, multiple sclerosis,

asthma, osteoarthritis, psoriasis, atherosclerosis, rhinitis, autoimmunity, and organ transplant rejection. Most preferably, the immune-related diseases and disorders are selected from the group consisting of rheumatoid arthritis, chronic inflammatory bowel disease, chronic inflammatory pelvic disease, multiple sclerosis, asthma, osteoarthritis, psoriasis, atherosclerosis, rhinitis, autoimmunity, and organ transplantation.

Substances useful for treatment of kinase-related disorders or diseases preferably show positive results in one or more *in vitro* assays for an activity corresponding to treatment of the disease or disorder in question (Examples of such assays are provided in the references in section VI, below; and in Example 7, herein). Examples of substances that can be screened for favorable activity are provided and referenced in section VI, below. The substances that modulate the activity of the kinases preferably include, but are not limited to, antisense oligonucleotides and inhibitors of protein kinases, as determined by methods and screens referenced in section VI and Example 7, below.

The term "preventing" refers to decreasing the probability that an organism contracts or develops an abnormal condition.

The term "treating" refers to having a therapeutic effect and at least partially alleviating or abrogating an abnormal condition in the organism.

The term "therapeutic effect" refers to the inhibition or activation factors causing or contributing to the abnormal condition. A therapeutic effect relieves to some extent one or more of the symptoms of the abnormal condition. In reference to the treatment of abnormal conditions, a therapeutic effect can refer to one or more of the following: (a) an increase in the proliferation, growth, and/or differentiation of cells; (b) inhibition (*i.e.*, slowing or stopping) of cell death; (c) inhibition of degeneration; (d) relieving to some extent one or more of the symptoms associated with the abnormal condition; and (e) enhancing the function of the affected population of cells. Compounds demonstrating efficacy against abnormal conditions can be identified as described herein.

The term "abnormal condition" refers to a function in the cells or tissues of an organism that deviates from their normal functions in that organism. An abnormal condition can relate to cell proliferation, cell differentiation, or cell survival.

Abnormal cell proliferative conditions include cancers such as fibrotic and mesangial disorders, abnormal angiogenesis and vasculogenesis, wound healing, psoriasis, diabetes mellitus, and inflammation.

Abnormal differentiation conditions include, but are not limited to neurodegenerative disorders, slow wound healing rates, and slow tissue grafting healing rates.

Abnormal cell survival conditions relate to conditions in which programmed cell death (apoptosis) pathways are activated or abrogated. A number of protein kinases are associated with the apoptosis pathways. Aberrations in the function of any one of the protein kinases could lead to cell immortality or premature cell death.

The term "aberration", in conjunction with the function of a kinase in a signal transduction process, refers to a kinase that is over- or under-expressed in an organism, mutated such that its catalytic activity is lower or higher than wild-type protein kinase activity, mutated such that it can no longer interact with a natural binding partner, is no longer modified by another protein kinase or protein phosphatase, or no longer interacts with a natural binding partner.

The term "administering" relates to a method of incorporating a compound into cells or tissues of an organism. The abnormal condition can be prevented or treated when the cells or tissues of the organism exist within the organism or outside of the organism. Cells existing outside the organism can be maintained or grown in cell culture dishes. For cells harbored within the organism, many techniques exist in the art to administer compounds, including (but not limited to) oral, parenteral, dermal, injection, and aerosol applications. For cells outside of the organism, multiple techniques exist in the art to administer the compounds, including (but not limited to) cell microinjection techniques, transformation techniques, and carrier techniques.

The abnormal condition can also be prevented or treated by administering a compound to a group of cells having an aberration in a signal transduction pathway to an organism. The effect of administering a compound on organism function can then be monitored. The organism is preferably a mouse, rat, rabbit, guinea pig, or goat, more preferably a monkey or ape, and most preferably a human.

In another aspect, the invention features methods for detection of a kinase polypeptide in a sample as a diagnostic tool for diseases or disorders, wherein the method comprises the steps of: (a) contacting the sample with a nucleic acid probe

which hybridizes under hybridization assay conditions to a nucleic acid target region of a kinase polypeptide having an amino acid sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24, said probe comprising the nucleic acid sequence encoding the polypeptide, fragments thereof, and the complements of the sequences and fragments; and (b) detecting the presence or amount of the probe:target region hybrid as an indication of the disease.

In preferred embodiments of the invention, the disease or disorder is selected from the group consisting of rheumatoid arthritis, arteriosclerosis, autoimmune disorders, organ transplantation, myocardial infarction, cardiomyopathies, stroke, renal failure, oxidative stress-related neurodegenerative disorders, and cancer.

The kinase "target region" is the nucleotide base sequence selected from the group consisting of those set forth in SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, and SEQ ID NO: 12, or the corresponding full-length sequences, a functional derivative thereof, or a fragment thereof, to which the nucleic acid probe will specifically hybridize. Specific hybridization indicates that in the presence of other nucleic acids the probe only hybridizes detectably with the kinase of the invention's target region. Putative target regions can be identified by methods well known in the art consisting of alignment and comparison of the most closely related sequences in the database.

In preferred embodiments the nucleic acid probe hybridizes to a kinase target region encoding at least 6, 12, 75, 90, 105, 120, 150, 200, 250, 300 or 350 contiguous amino acids of a sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24, or the corresponding full-length amino acid sequence, a portion of any of these sequences that retains functional activity, as described herein, or a functional derivative thereof. Hybridization conditions should be such that hybridization occurs only with the kinase genes in the presence of other nucleic acid molecules. Under stringent hybridization conditions only highly complementary nucleic acid sequences hybridize. Preferably, such conditions prevent

hybridization of nucleic acids having more than 1 or 2 mismatches out of 20 contiguous nucleotides. Such conditions are defined *supra*.

The diseases for which detection of kinase genes in a sample could be diagnostic include diseases in which kinase nucleic acid (DNA and/or RNA) is amplified in comparison to normal cells. By "amplification" is meant increased numbers of kinase DNA or RNA in a cell compared with normal cells. In normal cells, kinases are typically found as single copy genes. In selected diseases, the chromosomal location of the kinase genes may be amplified, resulting in multiple copies of the gene, or amplification. Gene amplification can lead to amplification of kinase RNA, or kinase RNA can be amplified in the absence of kinase DNA amplification.

"Amplification" as it refers to RNA can be the detectable presence of kinase RNA in cells, since in some normal cells there is no basal expression of kinase RNA. In other normal cells, a basal level of expression of kinase exists, therefore in these cases amplification is the detection of at least 1-2-fold, and preferably more, kinase RNA, compared to the basal level.

The diseases that could be diagnosed by detection of kinase nucleic acid in a sample preferably include cancers. The test samples suitable for nucleic acid probing methods of the present invention include, for example, cells or nucleic acid extracts of cells, or biological fluids. The samples used in the above-described methods will vary based on the assay format, the detection method and the nature of the tissues, cells or extracts to be assayed. Methods for preparing nucleic acid extracts of cells are well known in the art and can be readily adapted in order to obtain a sample that is compatible with the method utilized.

The invention also features a method for detection of a kinase polypeptide in a sample as a diagnostic tool for a disease or disorder, wherein the method comprises: (a) comparing a nucleic acid target region encoding the kinase polypeptide in a sample, where the kinase polypeptide has an amino acid sequence selected from the group consisting those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24, or one or more fragments thereof, with a control nucleic acid target region encoding the kinase polypeptide, or one or more fragments thereof; and (b) detecting differences in sequence or amount between the target region and the control target region, as an

indication of the disease or disorder. Preferably the disease is selected from the group consisting of cancers, immune-related diseases and disorders, cardiovascular disease, brain or neuronal-associated diseases, and metabolic disorders. More specifically these diseases include cancer of tissues, blood, or hematopoietic origin, particularly those involving breast, colon, lung, prostate, cervical, brain, ovarian, bladder, or kidney; central or peripheral nervous system diseases and conditions including migraine, pain, sexual dysfunction, mood disorders, attention disorders, cognition disorders, hypotension, and hypertension; psychotic and neurological disorders, including anxiety, schizophrenia, manic depression, delirium, dementia, severe mental retardation and dyskinesias, such as Huntington's disease or Tourette's Syndrome; neurodegenerative diseases including Alzheimer's, Parkinson's, Multiple sclerosis, and Amyotrophic lateral sclerosis; viral or non-viral infections caused by HIV-1, HIV-2 or other viral- or prion-agents or fungal- or bacterial- organisms; metabolic disorders including Diabetes and obesity and their related syndromes, among others; cardiovascular disorders including reperfusion restenosis, coronary thrombosis, clotting disorders, unregulated cell growth disorders, atherosclerosis; ocular disease including glaucoma, retinopathy, and macular degeneration; inflammatory disorders including rheumatoid arthritis, chronic inflammatory bowel disease, chronic inflammatory pelvic disease, multiple sclerosis, asthma, osteoarthritis, psoriasis, atherosclerosis, rhinitis, autoimmunity, and organ transplant rejection.

The term "comparing" as used herein refers to identifying discrepancies between the nucleic acid target region isolated from a sample, and the control nucleic acid target region. The discrepancies can be in the nucleotide sequences, *e.g.* insertions, deletions, or point mutations, or in the amount of a given nucleotide sequence. Methods to determine these discrepancies in sequences are well-known to one of ordinary skill in the art. The "control" nucleic acid target region refers to the sequence or amount of the sequence found in normal cells, *e.g.* cells that are not diseased as discussed previously.

The summary of the invention described above is not limiting and other features and advantages of the invention will be apparent from the following detailed description of the invention, and from the claims.

BRIEF DESCRIPTION OF THE FIGURES

Figures 1A-1L show the nucleotide sequences for human protein kinases oriented in a 5' to 3' direction (SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, and SEQ ID NO: 12).

Figures 2A-2D show the amino acid sequences for the human protein kinases encoded by SEQ ID No. 1-12 in the direction of translation (SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24). Some of the sequences encode predicted stop codons within the coding region, indicated by an 'x.'

DETAILED DESCRIPTION OF THE INVENTION

The invention provides, *inter alia,* protein kinase and kinase-like genes, as well as fragments thereof, which have been identified in genomic databases. In part, the invention provides nucleic acid molecules that are capable of encoding polypeptides having a kinase or kinase-like activity. By reference to Tables 1 though 8, below, genes of the invention can be better understood. The invention additionally provides a number of different embodiments, such as those described below.

Nucleic Acids

Associations of chromosomal localizations for mapped genes with amplicons implicated in cancer are based on literature searches (PubMed http://www.ncbi.nlm.nih.gov/entrez/query.fcgi), OMIM searches (Online Mendelian Inheritance in Man, http://www.ncbi.nlm.nih.gov/Omim/searchomim.html) and the comprehensive database of cancer amplicons maintained by Knuutila, et al. (Knuutila, et al., DNA copy number amplifications in human neoplasms. Review of comparative genomic hybridization studies. Am J Pathol 152:1107-1123, 1998. http://www.helsinki.fi/~lgl_www/CMG.html). For many of the mapped genes, the cytogenetic region from Knuutila is listed followed by the number of cases with documented amplification and the total number of cases studied. Thus for SGK187, the entry "non-small cell lung cancer (12q24.1-24.3; 2/50)" means that the chromosomal position has been associated with non-small cell lung cancer, at position

12q24.1-24.3, which encompasses the SGK087's position, and the amplification has been noted in 2 of the 50 samples studied.

For single nucleotide polymorphisms, an accession number (for example, ss2014963 for SGK137 is given if the SNP is documented in dbSNP (the database of single nucleotide polymorphisms) maintained at NCBI (http://www.ncbi.nlm.nih.gov/SNP/index.html). The accession number for SNP can be used to retrieve the full SNP-containing sequence from this site. Candidate SNPs without a dbSNP accession number were identified by inspection of Blastn outputs of the patent sequences vs cDNA and genomic databases as indicated, for example, in Tables 6 and 7, provided in Example 1.

Nucleic Acid Probes, Methods, and Kits for Detection of Kinases

The invention additionally provides nucleic acid probes and uses therefor. A nucleic acid probe of the present invention may be used to probe an appropriate chromosomal or cDNA library by usual hybridization methods to obtain other nucleic acid molecules of the present invention. A chromosomal DNA or cDNA library may be prepared from appropriate cells according to recognized methods in the art (*cf.* "Molecular Cloning: A Laboratory Manual", second edition, Cold Spring Harbor Laboratory, Sambrook, Fritsch, & Maniatis, eds., 1989).

In the alternative, chemical synthesis can be carried out in order to obtain nucleic acid probes having nucleotide sequences which correspond to N-terminal and C-terminal portions of the amino acid sequence of the polypeptide of interest. The synthesized nucleic acid probes may be used as primers in a polymerase chain reaction (PCR) carried out in accordance with recognized PCR techniques, essentially according to PCR Protocols, "A Guide to Methods and Applications", Academic Press, Michael, *et al.*, eds., 1990, utilizing the appropriate chromosomal or cDNA library to obtain the fragment of the present invention.

One skilled in the art can readily design such probes based on the sequence disclosed herein using methods of computer alignment and sequence analysis known in the art ("Molecular Cloning: A Laboratory Manual", 1989, *supra*). The hybridization probes of the present invention can be labeled by standard labeling techniques such as with a radiolabel, enzyme label, fluorescent label, biotin-avidin label, chemiluminescence, and the like. After hybridization, the probes may be visualized using known methods.

The nucleic acid probes of the present invention include RNA, as well as DNA probes, such probes being generated using techniques known in the art. The nucleic acid probe may be immobilized on a solid support. Examples of such solid supports include, but are not limited to, plastics such as polycarbonate, complex carbohydrates such as agarose and sepharose, and acrylic resins, such as polyacrylamide and latex beads. Techniques for coupling nucleic acid probes to such solid supports are well known in the art.

The test samples suitable for nucleic acid probing methods of the present invention include, for example, cells or nucleic acid extracts of cells, or biological fluids. The samples used in the above-described methods will vary based on the assay format, the detection method and the nature of the tissues, cells or extracts to be assayed. Methods for preparing nucleic acid extracts of cells are well known in the art and can be readily adapted in order to obtain a sample which is compatible with the method utilized.

One method of detecting the presence of nucleic acids of the invention in a sample comprises (a) contacting said sample with the above-described nucleic acid probe under conditions such that hybridization occurs, and (b) detecting the presence of said probe bound to said nucleic acid molecule. One skilled in the art would select the nucleic acid probe according to techniques known in the art as described above. Samples to be tested include but should not be limited to RNA samples of human tissue.

A kit for detecting the presence of nucleic acids of the invention in a sample comprises at least one container means having disposed therein the above-described nucleic acid probe. The kit may further comprise other containers comprising one or more of the following: wash reagents and reagents capable of detecting the presence of bound nucleic acid probe. Examples of detection reagents include, but are not limited to radiolabelled probes, enzymatic labeled probes (horseradish peroxidase, alkaline phosphatase), and affinity labeled probes (biotin, avidin, or steptavidin). Preferably, the kit further comprises instructions for use.

In detail, a compartmentalized kit includes any kit in which reagents are contained in separate containers. Such containers include small glass containers, plastic containers or strips of plastic or paper. Such containers allow the efficient transfer of reagents from one compartment to another compartment such that the samples and reagents are not cross-contaminated and the agents or solutions of each

container can be added in a quantitative fashion from one compartment to another. Such containers will include a container which will accept the test sample, a container which contains the probe or primers used in the assay, containers which contain wash reagents (such as phosphate buffered saline, Tris-buffers, and the like), and containers which contain the reagents used to detect the hybridized probe, bound antibody, amplified product, or the like. One skilled in the art will readily recognize that the nucleic acid probes described in the present invention can readily be incorporated into one of the established kit formats which are well known in the art.

CATEGORIZATION OF THE POLYPEPTIDES ACCORDING TO THE INVENTION

For a number of protein kinases of the invention, there is provided a classification of the protein class and family to which it belongs, a summary of non-catalytic protein motifs, as well as a chromosomal location. This information is useful in determing function, regulation and/or therapeutic utility for each of the proteins. Amplification of chromosomal region can be associated with various cancers. For amplicons discussed in this application, the source of information was Knuutila, et al (Knuutila S, Björkqvist A-M, Autio K, Tarkkanen M, Wolf M, Monni O, Szymanska J, Larramendy ML, Tapper J, Pere H, El-Rifai W, Hemmer S, Wasenius V-M, Vidgren V & Zhu Y: DNA copy number amplifications in human neoplasms. Review of comparative genomic hybridization studies. Am J Pathol 152:1107-1123, 1998. http://www.helsinki.fi/~lgl_www/CMG.html).

The kinase classification and protein domains often reflect pathways, cellular roles, or mechanisms of up- or down-stream regulation. Also disease-relevant genes often occur in families of related genes. For example, if one member of a kinase family functions as an oncogene, a tumor suppressor, or has been found to be disrupted in an immune, neurologic, cardiovascular, or metabolic disorder, frequently other family members may play a related role.

The expression analysis organizes kinases into groups that are transcriptionally upregulated in tumors and those that are more restricted to specific tumor types such as melanoma or prostate. This analysis also identifies genes that are regulated in a cell cycle dependent manner, and are therefore likely to be involved in maintaining cell cycle checkpoints, entry, progression, or exit from mitosis, oversee DNA repair, or are involved in cell proliferation and genome stability. Expression data also can identify genes expressed in endothelial sources or other tissues that

suggest a role in angiogenesis, thereby implicating them as targets for control of diseases that have an angiogenic component, such as cancer, endometriosis, retinopathy and macular degeneration, and various ischemic or vascular pathologies. A proteins' role in cell survival can also be suggested based on restricted expression in cells subjected to external stress such as oxidative damage, hypoxia, drugs such as cisplatinum, or irradiation. Metastases-associated genes can be implicated when expression is restricted to invading regions of a tumor, or is only seen in local or distant metastases compared to the primary tumor, or when a gene is upregulated during cell culture models of invasion, migration, or motility.

Chromosomal location can identify candidate targets for a tumor amplicon or a tumor-suppressor locus. Summaries of prevalent tumor amplicons are available in the literature, and can identify tumor types to experimentally be confirmed to contain amplified copies of a kinase gene which localizes to an adjacent region.

As described herein, the polypeptides of the present invention can be classified among several groups. The salient features related to the biological and clinical implications of these different groups are described hereafter in more general terms.

A more specific characterization of the polypeptides of the invention, including potential biological and clinical implications, is provided, *e.g.*, in EXAMPLES 2 and 5.

CLASSIFICATION OF POLYPEPTIDES EXHIBITING KINASE ACTIVITY

The following information also is referenced, for example, at Tables 1 and 2.

AGC Group

Family members are described that belong to the AGC group of protein kinases. The AGC group of protein kinases includes as its major prototypes protein kinase C (PKC), cAMP-dependent protein kinases (PKA), the G protein-coupled receptor kinases (ARK and rhodopsin kinase (GRK1)) as well as p70S6K and AKT.

Potential biological and clinical implications of the novel AGC group protein kinases are described in Example 6. Novel AGC group kinases include: SEQ ID NO: 13.

CAMK Group

Family members are described that belong to the CAMK group of protein kinases. The CAMK group of protein kinases includes as its major prototypes the calmodulin-dependent protein kinases, elongation factor-2 kinases, phosphorylase kinase and the Snf1 and cAMP-dependent family of protein kinases.

Potential biological and clinical implications of the novel CAMK group of protein kinases are described in Example 6. Novel CAMK group of protein kinases include: SEQ ID NO: 14.

Casein Kinase Group

Family members are described that belong to the casein kinase (CKI) group of protein kinases. The Casein kinase (CK) group of protein kinases includes as its major prototypes casein kinaseI (CKI) and casein kinaseII (CKII). Both CKI and CKII are ubiquitous, constitutively-active, second-messenger-independent kinases, highly conserved enzymes that exist in multiple isoforms. CKI functions in vesicular trafficking, DNA repair, cell cycle progression and cytokinesis (Cell Signal 1998 Nov;10(10):699-711). CK2 functions in cell cycle progression in non-neural cells. CK2 has also been implicated in multiple signaling pathways in normal and disease states of the mammalian nervous systems (Prog Neurobiol 2000 Feb;60(3):211-46). Potential biological and clinical implications of the novel casein kinase group of protein kinases are described in Example 6. Novel casein kinase protein kinases include: SEQ ID NO: 15.

"Other" group

Family members are described that belong to the "Other" group of protein kinases. Within this group of protein kinases are members that have recognizable catalytic motifs that are identifiable by a hidden Markov model analysis, but fail to cluster with other protein kinases on the basis of their amino acid sequence homology over the catalytic region.

Potential biological and clinical implications of the novel protein kinases belonging to the Other group are described in Example 6. Novel "Other" protein kinases include: SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, and SEQ ID NO: 20.

The TK group

Family members are described that belong to the tyrosine kinase (TK) group of protein kinases. The TK group of protein kinases includes as its major prototypes the cytoplasmic and receptor families of protein kinases. Proteins within this group include SEQ ID NO: 21 and SEQ ID NO: 22.

CLASSIFICATION OF POLYPEPTIDES EXHIBITING KINASE-LIKE ACTIVITY

Novel members also are described that belong to the protein kinase (PK)-like superfamily of protein kinases. The PK-like superfamily of kinases comprise enzymes that phosphorylate non-protein substrates (i.e. metabolites, lipids) such as amino acid kinases (i.e. choline kinase), nucleotides (i.e. guanylate kinases) and lipid kinases (i.e. phospho inositide kinases) , as decribed in the EXAMPLES and Tables.

Two genes in this patent: SGK119 (SEQ ID NO: 11, encoding SEQ ID NO: 23) and SGK387 (SEQ ID NO: 12, encoding SEQ ID NO: 24), are PK-like (*i.e.*, do not have PK domains). In table 3, in the columns listing the protein kinase domain boundaries for the other genes in this patent, these genes have the entry "PK-like", to indicate that these genes do not have cononical protein kinase domains. The boundaries of their catalytic domains – a guanyl cyclase domain for SGK119 (SEQ ID NO: 11) and a phosphoinositide kinase domain for SGK387 (SEQ ID NO: 12) - are listed in the column "Additional domains". The following paragraphs describe the hidden Markov profiles used to identify these non-kinase domains.

Protein kinase C terminal domain (PF00433) is approximately 49 amino acids long. It is built from 21 members and is found protein kinase C family members.

Armadillo/beta-catenin-like repeats (PF00514) are approximately 41 amino acids long. They are built from 256 members that include the adenomatous polyposis coli (APC) tumor suppressor and Beta catenin.

Leucine Rich Repeats (PF00560) are approximately 23 amino acids long. They are built from 2587 members that include numerous proteins with diverse functions and cellular locations. Proteins with leucine-rich repeats are typically involved in protein-protein interactions. Ribonuclease inhibitor is an example of a

protein that contains leucine rich repeats. It is a cytoplasmic protein that tightly binds and inhibits ribonucleases of the pancreatic ribonuclease superfamily. Ribonuclease inhibitor has been shown to play a role in the regulation of RNA and in angiogenesis.

The TBC domain (PF00566) is approximately 233 amino acids long. It is built from 11 members that include proteins like Tbc-1, the tre-2 oncogene and the yeast regulators of mitosis, BUB2 and cdc16. Tbc1 is a nuclear protein with tissue expression being cell- and stage-specific. Tbc1 may play a role in the cell cycle and differentiation of various tissues.

Rhodanese-like domain (PF00581) is approximately 111 amino acids long. It is built from 49 members. Rhodanese catalyzes the transfer of the sulfane atom of thiosulfate to cyanide, to form sulfite and thiocyanate. Rhodanese is found in vertebrates, where it is a mitochondrial enzyme involved in forming iron-sulfur complexes and cyanide detoxification. Other bacterial proteins closely related to rhodanese include rhdA (*Azotobacter vinelandii*), sseA (*Escherichia coli*), and *cysA* (*Saccharopolyspora erythraea*).

Ephrin receptor ligand binding domain (PF01404) is approximately 172 amino acids long. It is built from 15 members that include the Eph receptors, which bind a group of cell-membrane-anchored ligands known as ephrins. The Eph receptors make up the largest subfamily of receptor tyrosine kinases, being predominantly expressed in the developing and adult nervous system. The Eph receptors have also been shown to play an important role in contact-mediated axon guidance, axon fasciculation and cell migration.

Fibronectin type III domain (PF00041) is approximately 86 amino acids long. It is built from 108 members that include a receptor supergroup typically involved with adhesive functions and a number of receptors for lymphokines, hematopoeitic growth factors and growth hormone-related molecules.

SAM domain (Sterile alpha motif) (PF00536) is approximately 64 amino acids long. It is built from 108 members that include 40 EPH-related receptor tyrosine kinases (RPTK), Drosophila bicaudal-C, a p53 from Loligo forbesi, and diacyglycerol-kinase isoform delta. The SAM domain is an evolutionary conserved protein binding domain involved in the regulation of many developmental processes. The EPH related RPTKs contain a conserved tyrosine that is likely to mediate cell-cell initiated signal transduction via the binding of SH2 containing proteins to phosphotyrosine.

The ankyrin domain (PF00023) is approximately 33 amino acids long. It is built from 2220 members that include the ankyrin family of structural proteins, CDK inhibitors such as p19INK4d, and other signaling proteins such as the nuclear factor NF-kappa-b p50 subunit and Bcl3 (b-cell lymphoma 3-encoded protein) among others. The ankyrin repeats generally consist of a beta, alpha, alpha, beta order of secondary structures. The repeats associate to form a higher order structure.

Adenylate and Guanylate cyclase catalytic domain (PF00211) is approximately 161 amino acids long. It is built from 24 members that include the receptors for natriuretic peptides (ANF) including GC-A and GC-B. Both cytoplasmic and membrane proteins have been shown to contain adenylate and guanylate cyclase catalytic domain.

Phosphatidylinositol 3- and 4-kinases (PF00454) is approximately 208 amino acids long. It is built from 33 members that include PI3 kinase and PI4 kinase both of which are involved in phosphorylating phosphoinositides on the 3- and 4- hydroxyl group of the inositol ring respectively. This phosphorylation is believed to be involved in the formation of second messengers for cell signaling.

FATC domain (FRAP, ATM, TRRAP Cterminal) (PF02260) is approximately 32 amino acids long. It is built from 15 members that include PIK related kinases.

THERAPEUTIC METHODS ACCORDING TO THE INVENTION:

Diagnostics:

The invention provides methods for detecting a polypeptide in a sample as a diagnostic tool for diseases or disorders, wherein the method comprises the steps of: (a) contacting the sample with a nucleic acid probe which hybridizes under hybridization assay conditions to a nucleic acid target region of a polypeptide selected from the group consisting of SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24, said probe comprising the nucleic acid sequence encoding the polypeptide, fragments thereof, and the complements of the sequences and fragments; and (b) detecting the presence or amount of the probe:target region hybrid as an indication of the disease.

In preferred embodiments of the invention, the disease or disorder is selected from the group consisting of rheumatoid arthritis, atherosclerosis, autoimmune

disorders, organ transplantation, myocardial infarction, cardiomyopathies, stroke, renal failure, oxidative stress-related neurodegenerative disorders, metabolic disorder including diabetes, reproductive disorders including infertility, and cancer.

Hybridization conditions should be such that hybridization occurs only with the genes in the presence of other nucleic acid molecules. Under stringent hybridization conditions only highly complementary nucleic acid sequences hybridize. Preferably, such conditions prevent hybridization of nucleic acids having 1 or 2 mismatches out of 20 contiguous nucleotides. Such conditions are defined supra.

The diseases for which detection of genes in a sample could be diagnostic include diseases in which nucleic acid (DNA and/or RNA) is amplified in comparison to normal cells. By "amplification" is meant increased numbers of DNA or RNA in a cell compared with normal cells.

"Amplification" as it refers to RNA can be the detectable presence of RNA in cells, since in some normal cells there is no basal expression of RNA. In other normal cells, a basal level of expression exists, therefore in these cases amplification is the detection of at least 1-2-fold, and preferably more, compared to the basal level.

The diseases that could be diagnosed by detection of nucleic acid in a sample preferably include cancers. The test samples suitable for nucleic acid probing methods of the present invention include, for example, cells or nucleic acid extracts of cells, or biological fluids. The samples used in the above-described methods will vary based on the assay format, the detection method and the nature of the tissues, cells or extracts to be assayed. Methods for preparing nucleic acid extracts of cells are well known in the art and can be readily adapted in order to obtain a sample that is compatible with the method utilized.

Antibodies, Hybridomas, Methods of Use and Kits for Detection of Kinases

The present invention relates to an antibody having binding affinity to a kinase of the invention. The polypeptide may have the amino acid sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24, or a functional derivative thereof, or at least 9 contiguous amino acids thereof (preferably, at least 20, 30, 35, or 40 contiguous amino acids thereof).

The present invention also relates to an antibody having specific binding affinity to a kinase of the invention. Such an antibody may be isolated by comparing its binding affinity to a kinase of the invention with its binding affinity to other polypeptides. Those which bind selectively to a kinase of the invention would be chosen for use in methods requiring a distinction between a kinase of the invention and other polypeptides. Such methods could include, but should not be limited to, the analysis of altered kinase expression in tissue containing other polypeptides.

The kinases of the present invention can be used in a variety of procedures and methods, such as for the generation of antibodies, for use in identifying pharmaceutical compositions, and for studying DNA/protein interaction.

The kinases of the present invention can be used to produce antibodies or hybridomas. One skilled in the art will recognize that if an antibody is desired, such a peptide could be generated as described herein and used as an immunogen. The antibodies of the present invention include monoclonal and polyclonal antibodies, as well fragments of these antibodies, and humanized forms. Humanized forms of the antibodies of the present invention may be generated using one of the procedures known in the art such as chimerization or CDR grafting.

The present invention also relates to a hybridoma which produces the above-described monoclonal antibody, or binding fragment thereof. A hybridoma is an immortalized cell line which is capable of secreting a specific monoclonal antibody.

In general, techniques for preparing monoclonal antibodies and hybridomas are well known in the art (Campbell, "Monoclonal Antibody Technology: Laboratory Techniques in Biochemistry and Molecular Biology," Elsevier Science Publishers, Amsterdam, The Netherlands, 1984; St. Groth *et al., J. Immunol. Methods* 35:1-21, 1980). Any animal (mouse, rabbit, and the like) which is known to produce antibodies can be immunized with the selected polypeptide. Methods for immunization are well known in the art. Such methods include subcutaneous or intraperitoneal injection of the polypeptide. One skilled in the art will recognize that the amount of polypeptide used for immunization will vary based on the animal which is immunized, the antigenicity of the polypeptide and the site of injection.

The polypeptide may be modified or administered in an adjuvant in order to increase the peptide antigenicity. Methods of increasing the antigenicity of a polypeptide are well known in the art. Such procedures include coupling the antigen

with a heterologous protein (such as globulin or β-galactosidase) or through the inclusion of an adjuvant during immunization.

For monoclonal antibodies, spleen cells from the immunized animals are removed, fused with myeloma cells, such as SP2/0-Agl4 myeloma cells, and allowed to become monoclonal antibody producing hybridoma cells. Any one of a number of methods well known in the art can be used to identify the hybridoma cell which produces an antibody with the desired characteristics. These include screening the hybridomas with an ELISA assay, western blot analysis, or radioimmunoassay (Lutz *et al.*, *Exp. Cell Res.* 175:109-124, 1988). Hybridomas secreting the desired antibodies are cloned and the class and subclass are determined using procedures known in the art (Campbell, "Monoclonal Antibody Technology: Laboratory Techniques in Biochemistry and Molecular Biology", *supra*, 1984).

For polyclonal antibodies, antibody-containing antisera is isolated from the immunized animal and is screened for the presence of antibodies with the desired specificity using one of the above-described procedures. The above-described antibodies may be detectably labeled. Antibodies can be detectably labeled through the use of radioisotopes, affinity labels (such as biotin, avidin, and the like), enzymatic labels (such as horseradish peroxidase, alkaline phosphatase, and the like) fluorescent labels (such as FITC or rhodamine, and the like), paramagnetic atoms, and the like. Procedures for accomplishing such labeling are well-known in the art, for example, *see* Stemberger *et al.*, *J. Histochem. Cytochem.* 18:315, 1970; Bayer *et al.*, *Meth. Enzym.* 62:308, 1979; Engval *et al.*, *Immunol.* 109:129, 1972; Goding, *J. Immunol. Meth.* 13:215, 1976. The labeled antibodies of the present invention can be used for *in vitro*, *in vivo*, and *in situ* assays to identify cells or tissues which express a specific peptide.

The above-described antibodies may also be immobilized on a solid support. Examples of such solid supports include plastics such as polycarbonate, complex carbohydrates such as agarose and sepharose, acrylic resins such as polyacrylamide and latex beads. Techniques for coupling antibodies to such solid supports are well known in the art (Weir *et al.*, "Handbook of Experimental Immunology" 4th Ed., Blackwell Scientific Publications, Oxford, England, Chapter 10, 1986; Jacoby *et al.*, *Meth. Enzym.* 34, Academic Press, N.Y., 1974). The immobilized antibodies of the

present invention can be used for *in vitro*, *in vivo*, and *in situ* assays as well as in immunochromotography.

Furthermore, one skilled in the art can readily adapt currently available procedures, as well as the techniques, methods and kits disclosed herein with regard to antibodies, to generate peptides capable of binding to a specific peptide sequence in order to generate rationally designed antipeptide peptides (Hurby *et al.*, "Application of Synthetic Peptides: Antisense Peptides", In Synthetic Peptides, A User's Guide, W.H. Freeman, NY, pp. 289-307, 1992; Kaspczak *et al.*, *Biochemistry* 28:9230-9238, 1989).

Anti-peptide peptides can be generated by replacing the basic amino acid residues found in the peptide sequences of the kinases of the invention with acidic residues, while maintaining hydrophobic and uncharged polar groups. For example, lysine, arginine, and/or histidine residues are replaced with aspartic acid or glutamic acid and glutamic acid residues are replaced by lysine, arginine or histidine.

The present invention also encompasses a method of detecting a kinase polypeptide in a sample, comprising: (a) contacting the sample with an above-described antibody, under conditions such that immunocomplexes form, and (b) detecting the presence of said antibody bound to the polypeptide. In detail, the methods comprise incubating a test sample with one or more of the antibodies of the present invention and assaying whether the antibody binds to the test sample. Altered levels of a kinase of the invention in a sample as compared to normal levels may indicate disease.

Conditions for incubating an antibody with a test sample vary. Incubation conditions depend on the format employed in the assay, the detection methods employed, and the type and nature of the antibody used in the assay. One skilled in the art will recognize that any one of the commonly available immunological assay formats (such as radioimmunoassays, enzyme-linked immunosorbent assays, diffusion-based Ouchterlony, or rocket immunofluorescent assays) can readily be adapted to employ the antibodies of the present invention. Examples of such assays can be found in Chard ("An Introduction to Radioimmunoassay and Related Techniques" Elsevier Science Publishers, Amsterdam, The Netherlands, 1986), Bullock *et al.* ("Techniques in Immunocytochemistry," Academic Press, Orlando, FL Vol. 1, 1982; Vol. 2, 1983; Vol. 3, 1985), Tijssen ("Practice and Theory of Enzyme

Immunoassays: Laboratory Techniques in Biochemistry and Molecular Biology," Elsevier Science Publishers, Amsterdam, The Netherlands, 1985).

The immunological assay test samples of the present invention include cells, protein or membrane extracts of cells, or biological fluids such as blood, serum, plasma, or urine. The test samples used in the above-described method will vary based on the assay format, nature of the detection method and the tissues, cells or extracts used as the sample to be assayed. Methods for preparing protein extracts or membrane extracts of cells are well known in the art and can readily be adapted in order to obtain a sample which is testable with the system utilized.

A kit contains all the necessary reagents to carry out the previously described methods of detection. The kit may comprise: (i) a first container means containing an above-described antibody, and (ii) second container means containing a conjugate comprising a binding partner of the antibody and a label. In another preferred embodiment, the kit further comprises one or more other containers comprising one or more of the following: wash reagents and reagents capable of detecting the presence of bound antibodies.

Examples of detection reagents include, but are not limited to, labeled secondary antibodies, or in the alternative, if the primary antibody is labeled, the chromophoric, enzymatic, or antibody binding reagents which are capable of reacting with the labeled antibody. The compartmentalized kit may be as described above for nucleic acid probe kits. One skilled in the art will readily recognize that the antibodies described in the present invention can readily be incorporated into one of the established kit formats which are well known in the art.

Isolation of Compounds Capable of Interacting with Kinases

The present invention also relates to a method of detecting a compound capable of binding to a kinase of the invention comprising incubating the compound with a kinase of the invention and detecting the presence of the compound bound to the kinase. The compound may be present within a complex mixture, for example, serum, body fluid, or cell extracts.

The present invention also relates to a method of detecting an agonist or antagonist of kinase activity or kinase binding partner activity comprising incubating cells that produce a kinase of the invention in the presence of a compound and detecting changes in the level of kinase activity or kinase binding partner activity.

The compounds thus identified would produce a change in activity indicative of the presence of the compound. The compound may be present within a complex mixture, for example, serum, body fluid, or cell extracts. Once the compound is identified it can be isolated using techniques well known in the art.

Modulating polypeptide activity:

The invention additionally provides methods for treating a disease or abnormal condition by administering to a patient in need of such treatment a substance that modulates the activity of a polypeptide selected from the group consisting of SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24. Preferably, the disease is selected from the group consisting of rheumatoid arthritis, atherosclerosis, autoimmune disorders, organ transplantation, myocardial infarction, cardiomyopathies, stroke, renal failure, oxidative stress-related neurodegenerative disorders, metabolic and reproductive disorders, and cancer.

Substances useful for treatment of disorders or diseases preferably show positive results in one or more assays for an activity corresponding to treatment of the disease or disorder in question Substances that modulate the activity of the polypeptides preferably include, but are not limited to, antisense oligonucleotides and inhibitors of protein kinases.

The term "preventing" refers to decreasing the probability that an organism contracts or develops an abnormal condition.

The term "treating" refers to having a therapeutic effect and at least partially alleviating or abrogating an abnormal condition in the organism.

The term "therapeutic effect" refers to the inhibition or activation factors causing or contributing to the abnormal condition. A therapeutic effect relieves to some extent one or more of the symptoms of the abnormal condition. In reference to the treatment of abnormal conditions, a therapeutic effect can refer to one or more of the following: (a) an increase in the proliferation, growth, and/or differentiation of cells; (b) inhibition (, slowing or stopping) of cell death; (c) inhibition of degeneration; (d) relieving to some extent one or more of the symptoms associated with the abnormal condition; and (e) enhancing the function of the affected population

of cells. Compounds demonstrating efficacy against abnormal conditions can be identified as described herein.

The term "abnormal condition" refers to a function in the cells or tissues of an organism that deviates from their normal functions in that organism. An abnormal condition can relate to cell proliferation, cell differentiation or cell survival. An abnormal condition may also include irregularities in cell cycle progression, i.e., irregularities in normal cell cycle progression through mitosis and meiosis.

Abnormal cell proliferative conditions include cancers such as fibrotic and mesangial disorders, abnormal angiogenesis and vasculogenesis, wound healing, psoriasis, diabetes mellitus, and inflammation.

Abnormal differentiation conditions include, but are not limited to, neurodegenerative disorders, slow wound healing rates, and slow tissue grafting healing rates.

Abnormal cell survival conditions may also relate to conditions in which programmed cell death (apoptosis) pathways are activated or abrogated. A number of protein kinases are associated with the apoptosis pathways. Aberrations in the function of any one of the protein kinases could lead to cell immortality or premature cell death.

The term "aberration", in conjunction with the function of a kinase in a signal transduction process, refers to a kinase that is over- or under-expressed in an organism, mutated such that its catalytic activity is lower or higher than wild-type protein kinase activity, mutated such that it can no longer interact with a natural binding partner, is no longer modified by another protein kinase or protein phosphatase, or no longer interacts with a natural binding partner.

The term "administering" relates to a method of incorporating a compound into cells or tissues of an organism. The abnormal condition can be prevented or treated when the cells or tissues of the organism exist within the organism or outside of the organism. Cells existing outside the organism can be maintained or grown in cell culture dishes. For cells harbored within the organism, many techniques exist in the art to administer compounds, including (but not limited to) oral, parenteral, dermal, injection, and aerosol applications. For cells outside of the organism, multiple techniques exist in the art to administer the compounds, including (but not limited to) cell microinjection techniques, transformation techniques and carrier techniques.

The abnormal condition can also be prevented or treated by administering a compound to a group of cells having an aberration in a signal transduction pathway to an organism. The effect of administering a compound on organism function can then be monitored. The organism is preferably a mouse, rat, rabbit, guinea pig or goat, more preferably a monkey or ape, and most preferably a human.

The present invention also encompasses a method of agonizing (stimulating) or antagonizing kinase associated activity in a mammal comprising administering to said mammal an agonist or antagonist to a polypeptide selected from the group consisting of SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24 in an amount sufficient to effect said agonism or antagonism. A method of treating diseases in a mammal with an agonist or antagonist of the activity of one of the kinases of the invention comprising administering the agonist or antagonist to a mammal in an amount sufficient to agonize or antagonize kinase-associated functions is also encompassed in the present application.

In an effort to discover novel treatments for diseases, biomedical researchers and chemists have designed, synthesized, and tested molecules that inhibit the function of protein kinases. Some small organic molecules form a class of compounds that modulate the function of protein kinases. Examples of molecules that have been reported to inhibit the function of protein kinases include, but are not limited to, bis monocyclic, bicyclic or heterocyclic aryl compounds (PCT WO 92/20642, published November 26, 1992 by Maguire *et al.*), vinylene-azaindole derivatives (PCT WO 94/14808, published July 7, 1994 by Ballinari *et al.*), 1-cyclopropyl-4-pyridyl-quinolones (U.S. Patent No. 5,330,992), styryl compounds (U.S. Patent No. 5,217,999), styryl-substituted pyridyl compounds (U.S. Patent No. 5,302,606), certain quinazoline derivatives (EP Application No. 0 566 266 A1), seleoindoles and selenides (PCT WO 94/03427, published February 17, 1994 by Denny *et al.*), tricyclic polyhydroxylic compounds (PCT WO 92/21660, published December 10, 1992 by Dow), and benzylphosphonic acid compounds (PCT WO 91/15495, published October 17, 1991 by Dow *et al*), all of which are incorporated by reference herein, including any drawings.

Compounds that can traverse cell membranes and are resistant to acid hydrolysis are potentially advantageous as therapeutics as they can become highly

bioavailable after being administered orally to patients. However, many of these protein kinase inhibitors only weakly inhibit the function of protein kinases. In addition, many inhibit a variety of protein kinases and will therefore cause multiple side-effects as therapeutics for diseases.

Some indolinone compounds, however, form classes of acid resistant and membrane permeable organic molecules. WO 96/22976 (published August 1, 1996 by Ballinari *et al.*) describes hydrosoluble indolinone compounds that harbor tetralin, naphthalene, quinoline, and indole substituents fused to the oxindole ring. These bicyclic substituents are in turn substituted with polar moieties including hydroxylated alkyl, phosphate, and ether moieties. U.S. Patent Application Serial Nos. 08/702,232, filed August 23, 1996, entitled "Indolinone Combinatorial Libraries and Related Products and Methods for the Treatment of Disease" by Tang *et al.* (Lyon & Lyon Docket No. 221/187) and 08/485,323, filed June 7, 1995, entitled "Benzylidene-Z-Indoline Compounds for the Treatment of Disease" by Tang *et al.* (Lyon & Lyon Docket No. 223/298) and International Patent Publications WO 96/40116, published December 19, 1996 by Tang, *et al.*, and WO 96/22976, published August 1, 1996 by Ballinari *et al.*, all of which are incorporated herein by reference in their entirety, including any drawings, figures, or tables, describe indolinone chemical libraries of indolinone compounds harboring other bicyclic moieties as well as monocyclic moieties fused to the oxindole ring. Applications 08/702,232, filed August 23, 1996, entitled "Indolinone Combinatorial Libraries and Related Products and Methods for the Treatment of Disease" by Tang *et al.* (Lyon & Lyon Docket No. 221/187), 08/485,323, filed June 7, 1995, entitled "Benzylidene-Z-Indoline Compounds for the Treatment of Disease" by Tang *et al.* (Lyon & Lyon Docket No. 223/298), and WO 96/22976, published August 1, 1996 by Ballinari *et al.* teach methods of indolinone synthesis, methods of testing the biological activity of indolinone compounds in cells, and inhibition patterns of indolinone derivatives.

Other examples of substances capable of modulating kinase activity include, but are not limited to, tyrphostins, quinazolines, quinoxolines, and quinolines. The quinazolines, tyrphostins, quinolines, and quinoxolines referred to above include well known compounds such as those described in the literature. For example, representative publications describing quinazolines include Barker *et al.*, EPO Publication No. 0 520 722 A1; Jones *et al.*, U.S. Patent No. 4,447,608; Kabbe *et al.*, U.S. Patent No. 4,757,072; Kaul and Vougioukas, U.S. Patent No. 5,316,553;

Kreighbaum and Comer, U.S. Patent No. 4,343,940; Pegg and Wardleworth, EPO Publication No. 0 562 734 A1; Barker *et al.*, (1991) *Proc. of Am. Assoc. for Cancer Research* 32:327; Bertino, J.R., (1979) *Cancer Research* 3:293-304; Bertino, J.R., (1979) *Cancer Research* 9(2 part 1):293-304; Curtin *et al.*, (1986) *Br. J. Cancer* 53:361-368; Fernandes *et al.*, (1983) *Cancer Research* 43:1117-1123 ; Ferris *et al. J. Org. Chem.* 44(2):173-178; Fry *et al.*, (1994) *Science* 265:1093-1095; Jackman *et al.*, (1981) *Cancer Research* 51:5579-5586; Jones *et al. J. Med. Chem.* 29(6):1114-1118; Lee and Skibo, (1987) *Biochemistry* 26(23):7355-7362; Lemus *et al.*, (1989) *J. Org. Chem.* 54:3511-3518; Ley and Seng, (1975) *Synthesis* 1975:415-522; Maxwell *et al.*, (1991) *Magnetic Resonance in Medicine* 17:189-196 ; Mini *et al.*, (1985) *Cancer Research* 45:325-330; Phillips and Castle, J. (1980) *Heterocyclic Chem.* 17(19):1489-1596; Reece *et al.*, (1977) *Cancer Research* 47(11):2996-2999; Sculier *et al.*, (1986) *Cancer Immunol. and Immunother.* 23, A65; Sikora *et al.*, (1984) *Cancer Letters* 23:289-295; Sikora *et al.*, (1988) *Analytical Biochem.* 172:344-355; all of which are incorporated herein by reference in their entirety, including any drawings.

Quinoxaline is described in Kaul and Vougioukas, U.S. Patent No. 5,316,553, incorporated herein by reference in its entirety, including any drawings.

Quinolines are described in Dolle *et al.*, (1994) *J. Med. Chem.* 37:2627-2629; MaGuire, J. (1994) *Med. Chem.* 37:2129-2131; Burke *et al.*, (1993) *J. Med. Chem.* 36:425-432 ; and Burke *et al.* (1992) *BioOrganic Med. Chem. Letters* 2:1771-1774, all of which are incorporated by reference in their entirety, including any drawings.

Tyrphostins are described in Allen *et al.*, (1993) *Clin. Exp. Immunol.* 91:141-156; Anafi *et al.*, (1993) *Blood* 82:12, 3524-3529; Baker *et al.*, (1992) *J. Cell Sci.* 102:543-555; Bilder *et al.*, (1991) *Amer. Physiol. Soc.* pp. 6363-6143:C721-C730; Brunton *et al.*, (1992) *Proceedings of Amer. Assoc. Cancer Rsch.* 33:558; Bryckaert *et al.*, (1992) *Exp. Cell Research* 199:255-261; Dong *et al.*, (1993) *J. Leukocyte Biology* 53:53-60; Dong *et al.*, (1993) *J. Immunol.* 151(5):2717-2724; Gazit *et al.*, (1989) *J. Med. Chem.* 32, 2344-2352; Gazit *et al.*, (1993) *J. Med. Chem.* 36:3556-3564; Kaur *et al.*, (1994) *Anti-Cancer Drugs* 5:213-222; King *et al.*, (1991) *Biochem. J.* 275:413-418; Kuo *et al.*, (1993) *Cancer Letters* 74:197-202; Levitzki, A., (1992) *The FASEB J.* 6:3275-3282; Lyall *et al.*, (1989) *J. Biol. Chem.* 264:14503-14509; Peterson *et al.*, (1993) *The Prostate* 22:335-345; Pillemer *et al.*, (1992) *Int. J. Cancer* 50:80-85; Posner *et al.*, (1993) *Molecular Pharmacology* 45:673-683; Rendu *et al.*, (1992) *Biol. Pharmacology* 44(5):881-888; Sauro and Thomas, (1993) *Life Sciences* 53:371-376;

Sauro and Thomas, (1993) *J. Pharm. and Experimental Therapeutics* 267(3):1191125; Wolbring *et al.*, (1994) *J. Biol. Chem.* 269(36):22470-22472; and Yoneda *et al.*, (1991) *Cancer Research* 51:4430-4435; all of which are incorporated herein by reference in their entirety, including any drawings.

Other compounds that could be used as modulators include oxindolinones such as those described in U.S. patent application Serial No. 08/702,232 filed August 23, 1996, incorporated herein by reference in its entirety, including any drawings.

RECOMBINANT DNA TECHNOLOGY:

DNA Constructs Comprising a Kinase Nucleic Acid Molecule and Cells Containing These Constructs:

The present invention also relates to a recombinant DNA molecule comprising, 5' to 3', a promoter effective to initiate transcription in a host cell and the above-described nucleic acid molecules. In addition, the present invention relates to a recombinant DNA molecule comprising a vector and an above-described nucleic acid molecule. The present invention also relates to a nucleic acid molecule comprising a transcriptional region functional in a cell, a sequence complementary to an RNA sequence encoding an amino acid sequence corresponding to the above-described polypeptide, and a transcriptional termination region functional in said cell. The above-described molecules may be isolated and/or purified DNA molecules.

The present invention also relates to a cell or organism that contains an above-described nucleic acid molecule and thereby is capable of expressing a polypeptide. The polypeptide may be purified from cells which have been altered to express the polypeptide. A cell is said to be "altered to express a desired polypeptide" when the cell, through genetic manipulation; is made to produce a protein which it normally does not produce or which the cell normally produces at lower levels. One skilled in the art can readily adapt procedures for introducing and expressing either genomic, cDNA, or synthetic sequences into either eukaryotic or prokaryotic cells.

A nucleic acid molecule, such as DNA, is said to be "capable of expressing" a polypeptide if it contains nucleotide sequences which contain transcriptional and translational regulatory information and such sequences are "operably linked" to nucleotide sequences which encode the polypeptide. An operable linkage is a linkage in which the regulatory DNA sequences and the DNA sequence sought to be expressed are connected in such a way as to permit gene sequence expression. The

precise nature of the regulatory regions needed for gene sequence expression may vary from organism to organism, but shall in general include a promoter region which, in prokaryotes, contains both the promoter (which directs the initiation of RNA transcription) as well as the DNA sequences which, when transcribed into RNA, will signal synthesis initiation. Such regions will normally include those 5'-non-coding sequences involved with initiation of transcription and translation, such as the TATA box, capping sequence, CAAT sequence, and the like.

If desired, the non-coding region 3' to the sequence encoding a kinase of the invention may be obtained by the above-described methods. This region may be retained for its transcriptional termination regulatory sequences, such as termination and polyadenylation. Thus, by retaining the 3'-region naturally contiguous to the DNA sequence encoding a kinase of the invention, the transcriptional termination signals may be provided. Where the transcriptional termination signals are not satisfactorily functional in the expression host cell, then a 3' region functional in the host cell may be substituted.

Two DNA sequences (such as a promoter region sequence and a sequence encoding a kinase of the invention) are said to be operably linked if the nature of the linkage between the two DNA sequences does not (1) result in the introduction of a frame-shift mutation, (2) interfere with the ability of the promoter region sequence to direct the transcription of a gene sequence encoding a kinase of the invention, or (3) interfere with the ability of the gene sequence of a kinase of the invention to be transcribed by the promoter region sequence. Thus, a promoter region would be operably linked to a DNA sequence if the promoter were capable of effecting transcription of that DNA sequence. Thus, to express a gene encoding a kinase of the invention, transcriptional and translational signals recognized by an appropriate host are necessary.

The present invention encompasses the expression of a gene encoding a kinase of the invention (or a functional derivative thereof) in either prokaryotic or eukaryotic cells. Prokaryotic hosts are, generally, very efficient and convenient for the production of recombinant proteins and are, therefore, one type of preferred expression system for kinases of the invention. Prokaryotes most frequently are represented by various strains of *E. coli*. However, other microbial strains may also be used, including other bacterial strains.

In prokaryotic systems, plasmid vectors that contain replication sites and control sequences derived from a species compatible with the host may be used. Examples of suitable plasmid vectors may include pBR322, pUC118, pUC119 and the like; suitable phage or bacteriophage vectors may include λgt10, λgt11 and the like; and suitable virus vectors may include pMAM-neo, pKRC and the like. Preferably, the selected vector of the present invention has the capacity to replicate in the selected host cell.

Recognized prokaryotic hosts include bacteria such as *E. coli*, *Bacillus*, *Streptomyces*, *Pseudomonas*, *Salmonella*, *Serratia*, and the like. However, under such conditions, the polypeptide will not be glycosylated. The prokaryotic host must be compatible with the replicon and control sequences in the expression plasmid.

To express a kinase of the invention (or a functional derivative thereof) in a prokaryotic cell, it is necessary to operably link the sequence encoding the kinase of the invention to a functional prokaryotic promoter. Such promoters may be either constitutive or, more preferably, regulatable (*i.e.*, inducible or derepressible). Examples of constitutive promoters include the *int* promoter of bacteriophage λ, the *bla* promoter of the β-lactamase gene sequence of pBR322, and the *cat* promoter of the chloramphenicol acetyl transferase gene sequence of pPR325, and the like. Examples of inducible prokaryotic promoters include the major right and left promoters of bacteriophage λ (P_L and P_R), the *trp*, λ*recA*, acZ, λ*acI*, and *gal* promoters of *E. coli*, the α-amylase (Ulmanen *et al.*, *J. Bacteriol.* 162:176-182, 1985) and the ς-28-specific promoters of *B. subtilis* (Gilman *et al.*, *Gene Sequence* 32:11-20, 1984), the promoters of the bacteriophages of *Bacillus* (Gryczan, in: The Molecular Biology of the Bacilli, Academic Press, Inc., NY, 1982), and *Streptomyces* promoters (Ward *et al.*, *Mol. Gen. Genet.* 203:468-478, 1986). Prokaryotic promoters are reviewed by Glick (*Ind. Microbiot.* 1:277-282, 1987), Cenatiempo (*Biochimie* 68:505-516, 1986), and Gottesman (*Ann. Rev. Genet.* 18:415-442, 1984).

Proper expression in a prokaryotic cell also requires the presence of a ribosome-binding site upstream of the gene sequence-encoding sequence. Such ribosome-binding sites are disclosed, for example, by Gold *et al.* (*Ann. Rev. Microbiol.* 35:365-404, 1981). The selection of control sequences, expression vectors, transformation methods, and the like, are dependent on the type of host cell used to express the gene. As used herein, "cell", "cell line", and "cell culture" may be

used interchangeably and all such designations include progeny. Thus, the words "transformants" or "transformed cells" include the primary subject cell and cultures derived therefrom, without regard to the number of transfers. It is also understood that all progeny may not be precisely identical in DNA content, due to deliberate or inadvertent mutations. However, as defined, mutant progeny have the same functionality as that of the originally transformed cell.

Host cells which may be used in the expression systems of the present invention are not strictly limited, provided that they are suitable for use in the expression of the kinase polypeptide of interest. Suitable hosts may often include eukaryotic cells. Preferred eukaryotic hosts include, for example, yeast, fungi, insect cells, mammalian cells either *in vivo*, or in tissue culture. Mammalian cells which may be useful as hosts include HeLa cells, cells of fibroblast origin such as VERO or CHO-K1, or cells of lymphoid origin and their derivatives. Preferred mammalian host cells include SP2/0 and J558L, as well as neuroblastoma cell lines such as IMR 332, which may provide better capacities for correct post-translational processing.

In addition, plant cells are also available as hosts, and control sequences compatible with plant cells are available, such as the cauliflower mosaic virus 35S and 19S, and nopaline synthase promoter and polyadenylation signal sequences. Another preferred host is an insect cell, for example the *Drosophila* larvae. Using insect cells as hosts, the *Drosophila* alcohol dehydrogenase promoter can be used (Rubin, *Science* 240:1453-1459, 1988). Alternatively, baculovirus vectors can be engineered to express large amounts of kinases of the invention in insect cells (Jasny, *Science* 238:1653, 1987; Miller *et al.*, in: *Genetic Engineering*, Vol. 8, Plenum, Setlow *et al.*, eds., pp. 277-297, 1986).

Any of a series of yeast expression systems can be utilized which incorporate promoter and termination elements from the actively expressed sequences coding for glycolytic enzymes that are produced in large quantities when yeast are grown in mediums rich in glucose. Known glycolytic gene sequences can also provide very efficient transcriptional control signals. Yeast provides substantial advantages in that it can also carry out post-translational modifications. A number of recombinant DNA strategies exist utilizing strong promoter sequences and high copy number plasmids which can be utilized for production of the desired proteins in yeast. Yeast recognizes leader sequences on cloned mammalian genes and secretes peptides bearing leader

sequences (*i.e.*, pre-peptides). Several possible vector systems are available for the expression of kinases of the invention in a mammalian host.

A wide variety of transcriptional and translational regulatory sequences may be employed, depending upon the nature of the host. The transcriptional and translational regulatory signals may be derived from viral sources, such as adenovirus, bovine papilloma virus, cytomegalovirus, simian virus, or the like, where the regulatory signals are associated with a particular gene sequence which has a high level of expression. Alternatively, promoters from mammalian expression products, such as actin, collagen, myosin, and the like, may be employed. Transcriptional initiation regulatory signals may be selected which allow for repression or activation, so that expression of the gene sequences can be modulated. Of interest are regulatory signals which are temperature-sensitive so that by varying the temperature, expression can be repressed or initiated, or are subject to chemical (such as metabolite) regulation.

Expression of kinases of the invention in eukaryotic hosts requires the use of eukaryotic regulatory regions. Such regions will, in general, include a promoter region sufficient to direct the initiation of RNA synthesis. Preferred eukaryotic promoters include, for example, the promoter of the mouse metallothionein I gene sequence (Hamer *et al., J. Mol. Appl. Gen.* 1:273-288, 1982); the TK promoter of Herpes virus (McKnight, *Cell* 31:355-365, 1982); the SV40 early promoter (Benoist *et al., Nature* (London) 290:304-31, 1981); and the yeast gal4 gene sequence promoter (Johnston *et al., Proc. Natl. Acad. Sci.* (USA) 79:6971-6975, 1982; Silver *et al., Proc. Natl. Acad. Sci.* (USA) 81:5951-5955, 1984).

Translation of eukaryotic mRNA is initiated at the codon which encodes the first methionine. For this reason, it is preferable to ensure that the linkage between a eukaryotic promoter and a DNA sequence which encodes a kinase of the invention (or a functional derivative thereof) does not contain any intervening codons which are capable of encoding a methionine (*i.e.*, AUG). The presence of such codons results either in the formation of a fusion protein (if the AUG codon is in the same reading frame as the kinase of the invention coding sequence) or a frame-shift mutation (if the AUG codon is not in the same reading frame as the kinase of the invention coding sequence).

A nucleic acid molecule encoding a kinase of the invention and an operably linked promoter may be introduced into a recipient prokaryotic or eukaryotic cell

either as a nonreplicating DNA or RNA molecule, which may either be a linear molecule or, more preferably, a closed covalent circular molecule. Since such molecules are incapable of autonomous replication, the expression of the gene may occur through the transient expression of the introduced sequence. Alternatively, permanent expression may occur through the integration of the introduced DNA sequence into the host chromosome.

A vector may be employed which is capable of integrating the desired gene sequences into the host cell chromosome. Cells which have stably integrated the introduced DNA into their chromosomes can be selected by also introducing one or more markers which allow for selection of host cells which contain the expression vector. The marker may provide for prototrophy to an auxotrophic host, biocide resistance, *e.g.*, antibiotics, or heavy metals, such as copper, or the like. The selectable marker gene sequence can either be directly linked to the DNA gene sequences to be expressed, or introduced into the same cell by co-transfection. Additional elements may also be needed for optimal synthesis of mRNA. These elements may include splice signals, as well as transcription promoters, enhancers, and termination signals. cDNA expression vectors incorporating such elements include those described by Okayama (*Mol. Cell. Biol.* 3:280-289, 1983).

The introduced nucleic acid molecule can be incorporated into a plasmid or viral vector capable of autonomous replication in the recipient host. Any of a wide variety of vectors may be employed for this purpose. Factors of importance in selecting a particular plasmid or viral vector include: the ease with which recipient cells that contain the vector may be recognized and selected from those recipient cells which do not contain the vector; the number of copies of the vector which are desired in a particular host; and whether it is desirable to be able to "shuttle" the vector between host cells of different species.

Preferred prokaryotic vectors include plasmids such as those capable of replication in *E. coli* (such as, for example, pBR322, ColEl, pSC101, pACYC 184, πVX; "Molecular Cloning: A Laboratory Manual", 1989, *supra*). Bacillus plasmids include pC194, pC221, pT127, and the like (Gryczan, In: The Molecular Biology of the Bacilli, Academic Press, NY, pp. 307-329, 1982). Suitable *Streptomyces* plasmids include p1J101 (Kendall *et al.*, *J. Bacteriol.* 169:4177-4183, 1987), and streptomyces bacteriophages such as ϕC31 (Chater *et al.*, In: Sixth International Symposium on

Actinomycetales Biology, Akademiai Kaido, Budapest, Hungary, pp. 45-54, 1986). *Pseudomonas* plasmids are reviewed by John *et al.* (*Rev. Infect. Dis.* 8:693-704, 1986), and Izaki (*Jpn. J. Bacteriol.* 33:729-742, 1978).

Preferred eukaryotic plasmids include, for example, BPV, vaccinia, SV40, 2-micron circle, and the like, or their derivatives. Such plasmids are well known in the art (Botstein *et al.*, Miami Wntr. Symp. 19:265-274, 1982; Broach, In: "The Molecular Biology of the Yeast Saccharomyces: Life Cycle and Inheritance", Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, p. 445-470, 1981; Broach, *Cell* 28:203-204, 1982; Bollon *et al.*, *J. Clin. Hematol. Oncol.* 10:39-48, 1980; Maniatis, In: Cell Biology: A Comprehensive Treatise, Vol. 3, Gene Sequence Expression, Academic Press, NY, pp. 563-608, 1980).

Once the vector or nucleic acid molecule containing the construct(s) has been prepared for expression, the DNA construct(s) may be introduced into an appropriate host cell by any of a variety of suitable means, *i.e.*, transformation, transfection, conjugation, protoplast fusion, electroporation, particle gun technology, calcium phosphate-precipitation, direct microinjection, and the like. After the introduction of the vector, recipient cells are grown in a selective medium, which selects for the growth of vector-containing cells. Expression of the cloned gene(s) results in the production of a kinase of the invention, or fragments thereof. This can take place in the transformed cells as such, or following the induction of these cells to differentiate (for example, by administration of bromodeoxyuracil to neuroblastoma cells or the like). A variety of incubation conditions can be used to form the peptide of the present invention. The most preferred conditions are those which mimic physiological conditions.

Transgenic Animals:

A variety of methods are available for the production of transgenic animals associated with this invention. DNA can be injected into the pronucleus of a fertilized egg before fusion of the male and female pronuclei, or injected into the nucleus of an embryonic cell (*e.g.*, the nucleus of a two-cell embryo) following the initiation of cell division (Brinster *et al.*, *Proc. Nat. Acad. Sci.* USA 82:4438-4442, 1985). Embryos can be infected with viruses, especially retroviruses, modified to carry inorganic-ion receptor nucleotide sequences of the invention.

Pluripotent stem cells derived from the inner cell mass of the embryo and stabilized in culture can be manipulated in culture to incorporate nucleotide sequences of the invention. A transgenic animal can be produced from such cells through implantation into a blastocyst that is implanted into a foster mother and allowed to come to term. Animals suitable for transgenic experiments can be obtained from standard commercial sources such as Charles River (Wilmington, MA), Taconic (Germantown, NY), Harlan Sprague Dawley (Indianapolis, IN), etc.

The procedures for manipulation of the rodent embryo and for microinjection of DNA into the pronucleus of the zygote are well known to those of ordinary skill in the art (Hogan *et al.*, *supra*). Microinjection procedures for fish, amphibian eggs and birds are detailed in Houdebine and Chourrout (*Experientia* 47:897-905, 1991). Other procedures for introduction of DNA into tissues of animals are described in U.S. Patent No. 4,945,050 (Sanford *et al.*, July 30, 1990).

By way of example only, to prepare a transgenic mouse, female mice are induced to superovulate. Females are placed with males, and the mated females are sacrificed by CO_2 asphyxiation or cervical dislocation and embryos are recovered from excised oviducts. Surrounding cumulus cells are removed. Pronuclear embryos are then washed and stored until the time of injection. Randomly cycling adult female mice are paired with vasectomized males. Recipient females are mated at the same time as donor females. Embryos then are transferred surgically. The procedure for generating transgenic rats is similar to that of mice (Hammer *et al.*, *Cell* 63:1099-1112, 1990).

Methods for the culturing of embryonic stem (ES) cells and the subsequent production of transgenic animals by the introduction of DNA into ES cells using methods such as electroporation, calcium phosphate/DNA precipitation and direct injection also are well known to those of ordinary skill in the art (Teratocarcinomas and Embryonic Stem Cells, A Practical Approach, E.J. Robertson, ed., IRL Press, 1987).

In cases involving random gene integration, a clone containing the sequence(s) of the invention is co-transfected with a gene encoding resistance. Alternatively, the gene encoding neomycin resistance is physically linked to the sequence(s) of the invention. Transfection and isolation of desired clones are carried out by any one of several methods well known to those of ordinary skill in the art (E.J. Robertson, *supra*).

DNA molecules introduced into ES cells can also be integrated into the chromosome through the process of homologous recombina-tion (Capecchi, *Science* 244:1288-1292, 1989). Methods for positive selection of the recombination event (*i.e.*, neo resistance) and dual positive-negative selection (*i.e.*, neo resistance and gancyclovir resistance) and the subsequent identification of the desired clones by PCR have been described by Capecchi, *supra* and Joyner *et al.* (*Nature* 338:153-156, 1989), the teachings of which are incorporated herein in their entirety including any drawings. The final phase of the procedure is to inject targeted ES cells into blastocysts and to transfer the blastocysts into pseudopregnant females. The resulting chimeric animals are bred and the offspring are analyzed by Southern blotting to identify individuals that carry the transgene. Procedures for the production of non-rodent mammals and other animals have been discussed by others (Houdebine and Chourrout, *supra*; Pursel *et al.*, *Science* 244:1281-1288, 1989; and Simms *et al.*, *Bio/Technology* 6:179-183, 1988).

Thus, the invention provides transgenic, nonhuman mammals containing a transgene encoding a kinase of the invention or a gene affecting the expression of the kinase. Such transgenic nonhuman mammals are particularly useful as an *in vivo* test system for studying the effects of introduction of a kinase, or regulating the expression of a kinase (*i.e.*, through the introduction of additional genes, antisense nucleic acids, or ribozymes).

A "transgenic animal" is an animal having cells that contain DNA which has been artificially inserted into a cell, which DNA becomes part of the genome of the animal which develops from that cell. Preferred transgenic animals are primates, mice, rats, cows, pigs, horses, goats, sheep, dogs and cats. The transgenic DNA may encode human kinases. Native expression in an animal may be reduced by providing an amount of antisense RNA or DNA effective to reduce expression of the receptor.

Gene Therapy:

Kinases or their genetic sequences will also be useful in gene therapy (reviewed in Miller, *Nature* 357:455-460, 1992). Miller states that advances have resulted in practical approaches to human gene therapy that have demonstrated positive initial results. The basic science of gene therapy is described in Mulligan (*Science* 260:926-931, 1993).

In one preferred embodiment, an expression vector containing a kinase coding sequence is inserted into cells, the cells are grown *in vitro* and then infused in large numbers into patients. In another preferred embodi-ment, a DNA segment containing a promoter of choice (for example a strong promoter) is transferred into cells containing an endogenous gene encoding kinases of the invention in such a manner that the promoter segment enhances expression of the endogenous kinase gene (for example, the promoter segment is transferred to the cell such that it becomes directly linked to the endogenous kinase gene).

The gene therapy may involve the use of an adenovirus containing kinase cDNA targeted to a tumor, systemic kinase increase by implantation of engineered cells, injection with kinase-encoding virus, or injection of naked kinase DNA into appropriate tissues.

Target cell populations may be modified by introducing altered forms of one or more components of the protein complexes in order to modulate the activity of such complexes. For example, by reducing or inhibiting a complex component activity within target cells, an abnormal signal transduction event(s) leading to a condition may be decreased, inhibited, or reversed. Deletion or missense mutants of a component, that retain the ability to interact with other components of the protein complexes but cannot function in signal transduction, may be used to inhibit an abnormal, deleterious signal transduction event.

Expression vectors derived from viruses such as retroviruses, vaccinia virus, adenovirus, adeno-associ-ated virus, herpes viruses, several RNA viruses, or bovine papilloma virus, may be used for delivery of nucleotide sequences (*e.g.*, cDNA) encod-ing recom-binant kinase of the invention protein into the targeted cell population (*e.g.*, tumor cells). Methods which are well known to those skilled in the art can be used to construct recombinant viral vectors containing coding sequences (Maniatis *et al.*, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, N.Y., 1989; Ausubel *et al.*, Current Protocols in Molecular Biology, Greene Publishing Associates and Wiley Interscience, N.Y., 1989). Alternatively, recombinant nucleic acid molecules encoding protein sequences can be used as naked DNA or in a reconstituted system *e.g.*, liposomes or other lipid systems for delivery to target cells (*e.g.*, Felgner *et al.*, *Nature* 337:387-8, 1989). Several other methods for the direct transfer of plasmid DNA into cells exist for use in human gene therapy and

involve targeting the DNA to receptors on cells by complexing the plasmid DNA to proteins (Miller, *supra*).

In its simplest form, gene transfer can be performed by simply injecting minute amounts of DNA into the nucleus of a cell, through a process of microinjection (Capecchi, *Cell* 22:479-88, 1980). Once recombinant genes are introduced into a cell, they can be recognized by the cell's normal mechanisms for transcription and translation, and a gene product will be expressed. Other methods have also been attempted for introducing DNA into larger numbers of cells. These methods include: transfection, wherein DNA is precipitated with calcium phosphate and taken into cells by pinocytosis (Chen *et al.*, *Mol. Cell Biol.* 7:2745-52, 1987); electroporation, wherein cells are exposed to large voltage pulses to introduce holes into the membrane (Chu *et al.*, *Nucleic Acids Res.* 15:1311-26, 1987); lipofection/liposome fusion, wherein DNA is packaged into lipophilic vesicles which fuse with a target cell (Felgner *et al.*, *Proc. Natl. Acad. Sci. USA.* 84:7413-7417, 1987); and particle bombardment using DNA bound to small projectiles (Yang *et al.*, *Proc. Natl. Acad. Sci.* 87:9568-9572, 1990). Another method for introducing DNA into cells is to couple the DNA to chemically modified proteins.

It has also been shown that adenovirus proteins are capable of destabilizing endosomes and enhancing the uptake of DNA into cells. The admixture of adenovirus to solutions containing DNA complexes, or the binding of DNA to polylysine covalently attached to adenovirus using protein crosslinking agents substantially improves the uptake and expression of the recombinant gene (Curiel *et al.*, *Am. J. Respir. Cell. Mol. Biol.*, 6:247-52, 1992).

As used herein "gene transfer" means the process of introducing a foreign nucleic acid molecule into a cell. Gene transfer is commonly performed to enable the expres-sion of a particular product encoded by the gene. The product may include a protein, polypeptide, anti-sense DNA or RNA, or enzymatically active RNA. Gene transfer can be performed in cultured cells or by direct administration into animals. Generally gene transfer involves the process of nucleic acid contact with a target cell by non-specific or receptor mediated interactions, uptake of nucleic acid into the cell through the membrane or by endocytosis, and release of nucleic acid into the cyto-plasm from the plasma membrane or endosome. Expression may require, in addition, movement of the nucleic acid into the nucleus of the cell and binding to appropriate nuclear factors for transcription.

As used herein "gene therapy" is a form of gene transfer and is included within the definition of gene transfer as used herein and specifically refers to gene transfer to express a therapeutic product from a cell *in vivo* or *in vitro*. Gene transfer can be performed *ex vivo* on cells which are then transplanted into a patient, or can be performed by direct administration of the nucleic acid or nucleic acid-protein complex into the patient.

In another preferred embodiment, a vector having nucleic acid sequences encoding a kinase polypeptide is provided in which the nucleic acid sequence is expressed only in specific tissue. Methods of achieving tissue-specific gene expression are set forth in International Publication No. WO 93/09236, filed November 3, 1992 and published May 13, 1993.

In all of the preceding vectors set forth above, a further aspect of the invention is that the nucleic acid sequence contained in the vector may include additions, deletions or modifications to some or all of the sequence of the nucleic acid, as defined above.

In another preferred embodiment, a method of gene replacement is set forth. "Gene replacement" as used herein means supplying a nucleic acid sequence which is capable of being expressed *in vivo* in an animal and thereby providing or augmenting the function of an endogenous gene which is missing or defective in the animal.

PHARMACEUTICAL FORMULATIONS AND ROUTES OF ADMINISTRATION

The compounds described herein can be administered to a human patient per se, or in pharmaceutical compositions where it is mixed with other active ingredients, as in combination therapy, or suitable carriers or excipient(s). Techniques for formulation and administration of the compounds of the instant application may be found in "Remington's Pharmaceutical Sciences," Mack Publishing Co., Easton, PA, latest edition.

Routes Of Administration:

Suitable routes of administration may, for example, include oral, rectal, transmucosal, or intestinal administration; parenteral delivery, including intramuscular, subcutaneous, intravenous, intramedullary injections, as well as intrathecal, direct intraventricular, intraperitoneal, intranasal, or intraocular injections.

Alternately, one may administer the compound in a local rather than systemic manner, for example, via injection of the compound directly into a solid tumor, often in a depot or sustained release formulation.

Furthermore, one may administer the drug in a targeted drug delivery system, for example, in a liposome coated with tumor-specific antibody. The liposomes will be targeted to and taken up selectively by the tumor.

Composition/Formulation:

The pharmaceutical compositions of the present invention may be manufactured in a manner that is itself known, *e.g.*, by means of conventional mixing, dissolving, granulating, dragee-making, levigating, emulsifying, encapsulating, entrapping or lyophilizing processes.

Pharmaceutical compositions for use in accordance with the present invention thus may be formulated in conventional manner using one or more physiologically acceptable carriers comprising excipients and auxiliaries which facilitate processing of the active compounds into preparations which can be used pharmaceutically. Proper formulation is dependent upon the route of administration chosen.

For injection, the agents of the invention may be formulated in aqueous solutions, preferably in physiologically compatible buffers such as Hanks's solution, Ringer's solution, or physiological saline buffer. For transmucosal administration, penetrants appropriate to the barrier to be permeated are used in the formulation. Such penetrants are generally known in the art.

For oral administration, the compounds can be formulated readily by combining the active compounds with pharmaceutically acceptable carriers well known in the art. Such carriers enable the compounds of the invention to be formulated as tablets, pills, dragees, capsules, liquids, gels, syrups, slurries, suspensions and the like, for oral ingestion by a patient to be treated. Suitable carriers include excipients such as, fillers such as sugars, including lactose, sucrose, mannitol, or sorbitol; cellulose preparations such as, for example, maize starch, wheat starch, rice starch, potato starch, gelatin, gum tragacanth, methyl cellulose, hydroxypropylmethyl- cellulose, sodium carboxymethylcellulose, and/or polyvinylpyrrolidone (PVP). If desired, disintegrating agents may be added, such as the cross-linked polyvinyl pyrrolidone, agar, or alginic acid or a salt thereof such as sodium alginate.

Dragee cores are provided with suitable coatings. For this purpose, concentrated sugar solutions may be used, which may optionally contain gum arabic, talc, polyvinyl pyrrolidone, carbopol gel, polyethylene glycol, and/or titanium dioxide, lacquer solutions, and suitable organic solvents or solvent mixtures. Dyestuffs or pigments may be added to the tablets or dragee coatings for identification or to characterize different combinations of active compound doses.

Pharmaceutical preparations which can be used orally include push-fit capsules made of gelatin, as well as soft, sealed capsules made of gelatin and a plasticizer, such as glycerol or sorbitol. The push-fit capsules can contain the active ingredients in admixture with filler such as lactose, binders such as starches, and/or lubricants such as talc or magnesium stearate and, optionally, stabilizers. In soft capsules, the active compounds may be dissolved or suspended in suitable liquids, such as fatty oils, liquid paraffin, or liquid polyethylene glycols. In addition, stabilizers may be added. All formulations for oral administration should be in dosages suitable for such administration.

For buccal administration, the compositions may take the form of tablets or lozenges formulated in conventional manner.

For administration by inhalation, the compounds for use according to the present invention are conveniently delivered in the form of an aerosol spray presentation from pressurized packs or a nebuliser, with the use of a suitable propellant, *e.g.*, dichlorodifluoromethane, trichlorofluoromethane, dichlorotetrafluoroethane, carbon dioxide or other suitable gas. In the case of a pressurized aerosol the dosage unit may be determined by providing a valve to deliver a metered amount. Capsules and cartridges of *e.g.* gelatin for use in an inhaler or insufflator may be formulated containing a powder mix of the compound and a suitable powder base such as lactose or starch.

The compounds may be formulated for parenteral administration by injection, *e.g.*, by bolus injection or continuous infusion. Formulations for injection may be presented in unit dosage form, *e.g.*, in ampoules or in multi-dose containers, with an added preservative. The compositions may take such forms as suspensions, solutions or emulsions in oily or aqueous vehicles, and may contain formulatory agents such as suspending, stabilizing and/or dispersing agents.

Pharmaceutical formulations for parenteral administration include aqueous solutions of the active compounds in water-soluble form. Additionally, suspensions

of the active compounds may be prepared as appropriate oily injection suspensions. Suitable lipophilic solvents or vehicles include fatty oils such as sesame oil, or synthetic fatty acid esters, such as ethyl oleate or triglycerides, or liposomes. Aqueous injection suspensions may contain substances which increase the viscosity of the suspension, such as sodium carboxymethyl cellulose, sorbitol, or dextran. Optionally, the suspension may also contain suitable stabilizers or agents which increase the solubility of the compounds to allow for the preparation of highly concentrated solutions.

Alternatively, the active ingredient may be in powder form for constitution with a suitable vehicle, *e.g.*, sterile pyrogen-free water, before use.

The compounds may also be formulated in rectal compositions such as suppositories or retention enemas, *e.g.*, containing conventional suppository bases such as cocoa butter or other glycerides.

In addition to the formulations described previously, the compounds may also be formulated as a depot preparation. Such long acting formulations may be administered by implantation (for example subcutaneously or intramuscularly) or by intramuscular injection. Thus, for example, the compounds may be formulated with suitable polymeric or hydrophobic materials (for example as an emulsion in an acceptable oil) or ion exchange resins, or as sparingly soluble derivatives, for example, as a sparingly soluble salt.

A pharmaceutical carrier for the hydrophobic compounds of the invention is a cosolvent system comprising benzyl alcohol, a nonpolar surfactant, a water-miscible organic polymer, and an aqueous phase. The cosolvent system may be the VPD co-solvent system. VPD is a solution of 3% w/v benzyl alcohol, 8% w/v of the nonpolar surfactant polysorbate 80, and 65% w/v polyethylene glycol 300, made up to volume in absolute ethanol. The VPD co-solvent system (VPD:D5W) consists of VPD diluted 1:1 with a 5% dextrose in water solution. This co-solvent system dissolves hydrophobic compounds well, and itself produces low toxicity upon systemic administration. Naturally, the proportions of a co-solvent system may be varied considerably without destroying its solubility and toxicity characteristics. Furthermore, the identity of the co-solvent components may be varied: for example, other low-toxicity nonpolar surfactants may be used instead of polysorbate 80; the fraction size of polyethylene glycol may be varied; other biocompatible polymers may

replace polyethylene glycol, *e.g.* polyvinyl pyrrolidone; and other sugars or polysaccharides may substitute for dextrose.

Alternatively, other delivery systems for hydrophobic pharmaceutical compounds may be employed. Liposomes and emulsions are well known examples of delivery vehicles or carriers for hydrophobic drugs. Certain organic solvents such as dimethylsulfoxide also may be employed, although usually at the cost of greater toxicity. Additionally, the compounds may be delivered using a sustained-release system, such as semipermeable matrices of solid hydrophobic polymers containing the therapeutic agent. Various sustained-release materials have been established and are well known by those skilled in the art. Sustained-release capsules may, depending on their chemical nature, release the compounds for a few weeks up to over 100 days. Depending on the chemical nature and the biological stability of the therapeutic reagent, additional strategies for protein stabilization may be employed.

The pharmaceutical compositions also may comprise suitable solid or gel phase carriers or excipients. Examples of such carriers or excipients include but are not limited to calcium carbonate, calcium phosphate, various sugars, starches, cellulose derivatives, gelatin, and polymers such as polyethylene glycols.

Many of the tyrosine or serine/threonine kinase modulating compounds of the invention may be provided as salts with pharmaceutically compatible counterions. Pharmaceutically compatible salts may be formed with many acids, including but not limited to hydrochloric, sulfuric, acetic, lactic, tartaric, malic, succinic, etc. Salts tend to be more soluble in aqueous or other protonic solvents that are the corresponding free base forms.

Suitable Dosage Regimens:

Pharmaceutical compositions suitable for use in the present invention include compositions where the active ingredients are contained in an amount effective to achieve its intended purpose. More specifically, a therapeutically effective amount means an amount of compound effective to prevent, alleviate or ameliorate symptoms of disease or prolong the survival of the subject being treated. Determination of a therapeutically effective amount is well within the capability of those skilled in the art, especially in light of the detailed disclosure provided herein.

Methods of determining the dosages of compounds to be administered to a patient and modes of administering compounds to an organism are disclosed in U.S.

Application Serial No. 08/702,282, filed August 23, 1996 and International patent publication number WO 96/22976, published August 1 1996, both of which are incorporated herein by reference in their entirety, including any drawings, figures or tables. Those skilled in the art will appreciate that such descriptions are applicable to the present invention and can be easily adapted to it.

The proper dosage depends on various factors such as the type of disease being treated, the particular composition being used and the size and physiological condition of the patient. Therapeutically effective doses for the compounds described herein can be estimated initially from cell culture and animal models. For example, a dose can be formulated in animal models to achieve a circulating concentration range that initially takes into account the IC_{50} as determined in cell culture assays. The animal model data can be used to more accurately determine useful doses in humans.

For any compound used in the methods of the invention, the therapeutically effective dose can be estimated initially from cell culture assays. For example, a dose can be formulated in animal models to achieve a circulating concentration range that includes the IC_{50} as determined in cell culture (*i.e.*, the concentration of the test compound which achieves a half-maximal inhibition of the tyrosine or serine/threonine kinase activity). Such information can be used to more accurately determine useful doses in humans.

Toxicity and therapeutic efficacy of the compounds described herein can be determined by standard pharmaceutical procedures in cell cultures or experimental animals, *e.g.*, for determining the LD_{50} (the dose lethal to 50% of the population) and the ED_{50} (the dose therapeutically effective in 50% of the population). The dose ratio between toxic and therapeutic effects is the therapeutic index and it can be expressed as the ratio between LD_{50} and ED_{50}. Compounds which exhibit high therapeutic indices are preferred. The data obtained from these cell culture assays and animal studies can be used in formulating a range of dosage for use in human. The dosage of such compounds lies preferably within a range of circulating concentrations that include the ED_{50} with little or no toxicity. The dosage may vary within this range depending upon the dosage form employed and the route of administration utilized. The exact formulation, route of administration and dosage can be chosen by the individual physician in view of the patient's condition. (See *e.g.*, Fingl *et al.*, 1975, in "The Pharmacological Basis of Therapeutics", Ch. 1 p.1).

In another example, toxicity studies can be carried out by measuring the blood cell composition. For example, toxicity studies can be carried out in a suitable animal model as follows: 1) the compound is administered to mice (an untreated control mouse should also be used); 2) blood samples are periodically obtained via the tail vein from one mouse in each treatment group; and 3) the samples are analyzed for red and white blood cell counts, blood cell composition and the percent of lymphocytes versus polymorphonuclear cells. A comparison of results for each dosing regime with the controls indicates if toxicity is present.

At the termination of each toxicity study, further studies can be carried out by sacrificing the animals (preferably, in accordance with the American Veterinary Medical Association guidelines Report of the American Veterinary Medical Assoc. Panel on Euthanasia:229-249, 1993). Representative animals from each treatment group can then be examined by gross necropsy for immediate evidence of metastasis, unusual illness or toxicity. Gross abnormalities in tissue are noted and tissues are examined histologically. Compounds causing a reduction in body weight or blood components are less preferred, as are compounds having an adverse effect on major organs. In general, the greater the adverse effect the less preferred the compound.

For the treatment of cancers the expected daily dose of a hydrophobic pharmaceutical agent is between 1 to 500 mg/day, preferably 1 to 250 mg/day, and most preferably 1 to 50 mg/day. Drugs can be delivered less frequently provided plasma levels of the active moiety are sufficient to maintain therapeutic effectiveness.

Plasma levels should reflect the potency of the drug. Generally, the more potent the compound the lower the plasma levels necessary to achieve efficacy.

Plasma half-life and biodistribution of the drug and metabolites in the plasma, tumors and major organs can also be determined to facilitate the selection of drugs most appropriate to inhibit a disorder. Such measurements can be carried out. For example, HPLC analysis can be performed on the plasma of animals treated with the drug and the location of radiolabeled compounds can be determined using detection methods such as X-ray, CAT scan and MRI. Compounds that show potent inhibitory activity in the screening assays, but have poor pharmacokinetic characteristics, can be optimized by altering the chemical structure and retesting. In this regard, compounds displaying good pharmacokinetic characteristics can be used as a model.

Dosage amount and interval may be adjusted individually to provide plasma levels of the active moiety which are sufficient to maintain the kinase modulating

effects, or minimal effective concentration (MEC). The MEC will vary for each compound but can be estimated from *in vitro* data; *e.g.*, the concentration necessary to achieve 50-90% inhibition of the kinase using the assays described herein. Dosages necessary to achieve the MEC will depend on individual characteristics and route of administration. However, HPLC assays or bioassays can be used to determine plasma concentrations.

Dosage intervals can also be determined using MEC value. Compounds should be administered using a regimen which maintains plasma levels above the MEC for 10-90% of the time, preferably between 30-90% and most preferably between 50-90%.

In cases of local administration or selective uptake, the effective local concentration of the drug may not be related to plasma concentration.

The amount of composition administered will, of course, be dependent on the subject being treated, on the subject's weight, the severity of the affliction, the manner of administration and the judgment of the prescribing physician.

Packaging:

The compositions may, if desired, be presented in a pack or dispenser device which may contain one or more unit dosage forms containing the active ingredient. The pack may for example comprise metal or plastic foil, such as a blister pack. The pack or dispenser device may be accompanied by instructions for administration. The pack or dispenser may also be accompanied with a notice associated with the container in form prescribed by a governmental agency regulating the manufacture, use, or sale of pharmaceuticals, which notice is reflective of approval by the agency of the form of the polynucleotide for human or veterinary administration. Such notice, for example, may be the labeling approved by the U.S. Food and Drug Administration for prescription drugs, or the approved product insert. Compositions comprising a compound of the invention formulated in a compatible pharmaceutical carrier may also be prepared, placed in an appropriate container, and labeled for treatment of an indicated condition. Suitable conditions indicated on the label may include treatment of a tumor, inhibition of angiogenesis, treatment of fibrosis, diabetes, and the like.

FUNCTIONAL DERIVATIVES

Also provided herein are functional derivatives of a polypeptide or nucleic acid of the invention. By "functional derivative" is meant a "chemical derivative," "fragment," or "variant," of the polypeptide or nucleic acid of the invention, which terms are defined below. A functional derivative retains at least a portion of the function of the protein, for example reactivity with an antibody specific for the protein, enzymatic activity or binding activity mediated through noncatalytic domains, which permits its utility in accordance with the present invention. It is well known in the art that due to the degeneracy of the genetic code numerous different nucleic acid sequences can code for the same amino acid sequence. Equally, it is also well known in the art that conservative changes in amino acid can be made to arrive at a protein or polypeptide that retains the functionality of the original. In both cases, all permutations are intended to be covered by this disclosure.

Included within the scope of this invention are the functional equivalents of the herein-described isolated nucleic acid molecules. The degeneracy of the genetic code permits substitution of certain codons by other codons that specify the same amino acid and hence would give rise to the same protein. The nucleic acid sequence can vary substantially since, with the exception of methionine and tryptophan, the known amino acids can be coded for by more than one codon. Thus, portions or all of the genes of the invention could be synthesized to give a nucleic acid sequence significantly different from one selected from the group consisting of those set forth in SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, and SEQ ID NO: 12. The encoded amino acid sequence thereof would, however, be preserved.

In addition, the nucleic acid sequence may comprise a nucleotide sequence which results from the addition, deletion or substitution of at least one nucleotide to the 5'-end and/or the 3'-end of the nucleic acid formula selected from the group consisting of those set forth in SEQ ID NO: 1, SEQ ID NO: 2, SEQ ID NO: 3, SEQ ID NO: 4, SEQ ID NO: 5, SEQ ID NO: 6, SEQ ID NO: 7, SEQ ID NO: 8, SEQ ID NO: 9, SEQ ID NO: 10, SEQ ID NO: 11, and SEQ ID NO: 12, or a derivative thereof. Any nucleotide or polynucleotide may be used in this regard, provided that its addition, deletion or substitution does not alter the amino acid sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID

NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24, which is encoded by the nucleotide sequence. For example, the present invention is intended to include any nucleic acid sequence resulting from the addition of ATG as an initiation codon at the 5'-end of the inventive nucleic acid sequence or its derivative, or from the addition of TTA, TAG or TGA as a termination codon at the 3'-end of the inventive nucleotide sequence or its derivative. Moreover, the nucleic acid molecule of the present invention may, as necessary, have restriction endonuclease recognition sites added to its 5'-end and/or 3'-end.

Such functional alterations of a given nucleic acid sequence afford an opportunity to promote secretion and/or processing of heterologous proteins encoded by foreign nucleic acid sequences fused thereto. All variations of the nucleotide sequence of the kinase genes of the invention and fragments thereof permitted by the genetic code are, therefore, included in this invention.

Further, it is possible to delete codons or to substitute one or more codons with codons other than degenerate codons to produce a structurally modified polypeptide, but one which has substantially the same utility or activity as the polypeptide produced by the unmodified nucleic acid molecule. As recognized in the art, the two polypeptides are functionally equivalent, as are the two nucleic acid molecules that give rise to their production, even though the differences between the nucleic acid molecules are not related to the degeneracy of the genetic code.

A "chemical derivative" of the complex contains additional chemical moieties not normally a part of the protein. Covalent modifications of the protein or peptides are included within the scope of this invention. Such modifications may be introduced into the molecule by reacting targeted amino acid residues of the peptide with an organic derivatizing agent that is capable of reacting with selected side chains or terminal residues, as described below.

Cysteinyl residues most commonly are reacted with alpha-haloacetates (and corresponding amines), such as chloroacetic acid or chloroacetamide, to give carboxymethyl or carboxyamidomethyl derivatives. Cysteinyl residues also are derivatized by reaction with bromotrifluoroacetone, chloroacetyl phosphate, N-alkylmaleimides, 3-nitro-2-pyridyl disulfide, methyl 2-pyridyl disulfide, p-chloromercuribenzoate, 2-chloromercuri-4-nitrophenol, or chloro-7-nitrobenzo-2-oxa-1,3-diazole.

Histidyl residues are derivatized by reaction with diethylprocarbonate at pH 5.5-7.0 because this agent is relatively specific for the histidyl side chain. Para-bromophenacyl bromide also is useful; the reaction is preferably performed in 0.1 M sodium cacodylate at pH 6.0.

Lysinyl and amino terminal residues are reacted with succinic or other carboxylic acid anhydrides. Derivatization with these agents has the effect or reversing the charge of the lysinyl residues. Other suitable reagents for derivatizing primary amine containing residues include imidoesters such as methyl picolinimidate; pyridoxal phosphate; pyridoxal; chloroborohydride; trinitrobenzenesulfonic acid; O-methylisourea; 2,4 pentanedione; and transaminase-catalyzed reaction with glyoxylate.

Arginyl residues are modified by reaction with one or several conventional reagents, among them phenylglyoxal, 2,3-butanedione, 1,2-cyclohexanedione, and ninhydrin. Derivatization of arginine residues requires that the reaction be performed in alkaline conditions because of the high pK_a of the guanidine functional group. Furthermore, these reagents may react with the groups of lysine as well as the arginine alpha-amino group.

Tyrosyl residues are well-known targets of modification for introduction of spectral labels by reaction with aromatic diazonium compounds or tetranitromethane. Most commonly, N-acetylimidizol and tetranitromethane are used to form O-acetyl tyrosyl species and 3-nitro derivatives, respectively.

Carboxyl side groups (aspartyl or glutamyl) are selectively modified by reaction with carbodiimide (R'-N-C-N-R') such as 1-cyclohexyl-3-(2-morpholinyl(4-ethyl) carbodiimide or 1-ethyl-3-(4-azonia-4,4-dimethylpentyl) carbodiimide. Furthermore, aspartyl and glutamyl residues are converted to asparaginyl and glutaminyl residues by reaction with ammonium ions.

Glutaminyl and asparaginyl residues are frequently deamidated to the corresponding glutamyl and aspartyl residues. Alternatively, these residues are deamidated under mildly acidic conditions. Either form of these residues falls within the scope of this invention.

Derivatization with bifunctional agents is useful, for example, for cross-linking the component peptides of the protein to each other or to other proteins in a complex to a water-insoluble support matrix or to other macromolecular carriers. Commonly used cross-linking agents include, for example, 1,1-bis(diazoacetyl)-2-

phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, for example, esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such as 3,3'-dithiobis(succinimidylpropionate), and bifunctional maleimides such as bis-N-maleimido-1,8-octane. Derivatizing agents such as methyl-3-[p-azidophenyl) dithiolpropioimidate yield photoactivatable intermediates that are capable of forming crosslinks in the presence of light. Alternatively, reactive water-insoluble matrices such as cyanogen bromide-activated carbohydrates and the reactive substrates described in U.S. Patent Nos. 3,969,287; 3,691,016; 4,195,128; 4,247,642; 4,229,537; and 4,330,440 are employed for protein immobilization.

Other modifications include hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, methylation of the alpha-amino groups of lysine, arginine, and histidine side chains (Creighton, T.E., Proteins: Structure and Molecular Properties, W.H. Freeman & Co., San Francisco, pp. 79-86 (1983)), acetylation of the N-terminal amine, and, in some instances, amidation of the C-terminal carboxyl groups.

Such derivatized moieties may improve the stability, solubility, absorption, biological half life, and the like. The moieties may alternatively eliminate or attenuate any undesirable side effect of the protein complex and the like. Moieties capable of mediating such effects are disclosed, for example, in Remington's Pharmaceutical Sciences, 18th ed., Mack Publishing Co., Easton, PA (1990).

The term "fragment" is used to indicate a polypeptide derived from the amino acid sequence of the proteins, of the complexes having a length less than the full-length polypeptide from which it has been derived. Such a fragment may, for example, be produced by proteolytic cleavage of the full-length protein. Preferably, the fragment is obtained recombinantly by appropriately modifying the DNA sequence encoding the proteins to delete one or more amino acids at one or more sites of the C-terminus, N-terminus, and/or within the native sequence. Fragments of a protein are useful for screening for substances that act to modulate signal transduction, as described herein. It is understood that such fragments may retain one or more characterizing portions of the native complex. Examples of such retained characteristics include: catalytic activity; substrate specificity; interaction with other molecules in the intact cell; regulatory functions; or binding with an antibody specific for the native complex, or an epitope thereof.

Another functional derivative intended to be within the scope of the present invention is a "variant" polypeptide which either lacks one or more amino acids or contains additional or substituted amino acids relative to the native polypeptide. The variant may be derived from a naturally occurring complex component by appropriately modifying the protein DNA coding sequence to add, remove, and/or to modify codons for one or more amino acids at one or more sites of the C-terminus, N-terminus, and/or within the native sequence. It is understood that such variants having added, substituted and/or additional amino acids retain one or more characterizing portions of the native protein, as described above.

A functional derivative of a protein with deleted, inserted and/or substituted amino acid residues may be prepared using standard techniques well-known to those of ordinary skill in the art. For example, the modified components of the functional derivatives may be produced using site-directed mutagenesis techniques (as exemplified by Adelman *et al.*, 1983, *DNA* 2:183) wherein nucleotides in the DNA coding the sequence are modified such that a modified coding sequence is modified, and thereafter expressing this recombinant DNA in a prokaryotic or eukaryotic host cell, using techniques such as those described above. Alternatively, proteins with amino acid deletions, insertions and/or substitutions may be conveniently prepared by direct chemical synthesis, using methods well-known in the art. The functional derivatives of the proteins typically exhibit the same qualitative biological activity as the native proteins.

TABLES AND DESCRIPTION THEREOF

Table 1 documents the name of each gene, the classification of each gene, the positions of the open reading frames within the sequence, and the length of the corresponding peptide. From left to right the data presented is as follows: "Gene Name", "ID#na", "ID#aa", "FL/Cat", "Superfamily", "Group", "Family", "NA_length", "ORF Start", "ORF End", "ORF Length", and "AA_length". "Gene name" refers to name given the sequence encoding the kinase or kinase-like enzyme. Each gene is represented by "SGK" designation followed by a number. The SGK name usually represents multiple overlapping sequences built into a single contiguous sequence (a "contig"). The "ID#na" and "ID#aa" refer to the identification numbers given each nucleic acid and amino acid sequence in this patent. "FL/Cat" refers to the length of the gene, with FL indicating full length, and "Cat' indicating that only the

catalytic domain is presented. "Partial" in this column indicates that the sequence encodes a partial protein kinase catalytic domain. "Superfamily" identifies whether the gene is a protein kinase or protein-kinase-like. "Group" and "Family" refer to the protein kinase classification defined by sequence homology and based on previously established phylogenetic analysis [Hardie, G. and Hanks S. The Protein Kinase Book, Academic Press (1995) and Hunter T. and Plowman, G. Trends in Biochemical Sciences (1977) 22:18-22 and Plowman G.D. et al. (1999) Proc. Natl. Acad. Sci. 96:13603-13610)]. "NA_length" refers to the length in nucleotides of the corresponding nucleic acid sequence. "ORF start" refers to the beginning nucleotide of the open reading frame. "ORF end" refers to the last nucleotide of the open reading frame, excluding the stop codon. "ORF length" refers to the length in nucleotides of the open reading frame (excluding the stop codon). "AA length" refers to the length in amino acids of the peptide encoded in the corresponding nuclei acid sequence.

Table 1

Open Reading Frames

Gene Name	ID#na	ID#aa	FL/Cat	Superfamily	Group	Family	NA_length	ORF Start	ORF End	ORF Length	AA_length
SGK216	1	13	partial	Protein Kinase	AGC	S6K	843	1	843	843	281
SGK237	2	14	Cat	Protein Kinase	CAMK	EMK	2562	1	2562	2562	854
SGK248	3	15	FL	Protein Kinase	CKI	CKI	1269	1	1269	1266	422
SGK223	4	16	FL	Protein Kinase	Other	AUR	4189	1	3732	3732	1244
SGK269	5	17	FL	Protein Kinase	Other	AUR	5241	1	5238	5238	1746
SGK258	6	18	FL	Protein kinase	Other	MLK	6042	1	6042	6042	2014
SGK382	7	19	FL	Protein Kinase	Other	Unique	3915	358	3038	2679	893
SGK424	8	20	FL	Protein kinase	Other	Unique	1683	1	1680	1680	560
SGK251	9	21	FL	Protein Kinase	TK	RTK-11	3903	280	3387	3108	1036
SGK307	10	22	FL	Protein kinase	TK	Unique	4739	123	4408	4284	1428
SGK119	11	23	partial	PK-like	GCyc	GCyc	819	1	819	819	273
SGK387	12	24	FL	PK-like	Inositol kinase	PI3K	5736	1	5733	5733	1911

Table 2 lists the following features of the genes described in this application: chromosomal localization, single nucleotide polymorphisms (SNPs), representation in dbEST, and repeat regions. From left to right the data presented is as follows: "Gene Name", "ID#na", "ID#aa", "FL/Cat", "Superfamily", "Group", "Family", "Chromosome", "SNPs", "dbEST_hits", & "Repeats". The contents of the first 7 columns (i.e. "Gene Name", "ID#na", "ID#aa", "FL/Cat", "Superfamily", "Group" and "Family") are as described above for Table 1. "Chromosome" refers to the cytogenetic localization of the gene. Information in the "SNPs" column describes the nucleic acid position and degenerate nature of candidate single nucleotide polymorphisms (SNPs). For example, for SGK307, the "SNPs" column contains "2460=R (aagagagtttacaR) ss1928275", indicating that there are instances of both an A and a G (R = G or A) at position 2640. The sequence preceding the SNP and the accession number from dbSNP are also given. "dbESThits" lists accession numbers of entries in the public database of ESTs (dbEST, http://www.ncbi.nlm.nih.gov/dbEST/index.html) that contain at least 150 bp of 100% identity to the corresponding gene. These ESTs were identified by blastn of dbEST. "Repeats" contains information about the location of short sequences, approximately 20 bp in length, that are of low complexity and that are present in several distinct genes. These repeats were identified by blastn of the DNA sequence against the non-redundant nucleic acid database at NCBI (nrna) (http://www.ncbi.nlm.nih.gov/SNP/snpblastByChr.html). To be included in this repeat column, the sequence could have 100% identity over its length and typically is present in at least 5 different genes.

Table 2
CHR, SNPs, dbEST, Repeats

ene Nam	ID#na	ID#aa	FL/Ca	Superfamily	Group	Family	Chromosome	SNPs	dbEST hits	Repeats
SGK216	1	13	partial	Protein Kinas	AGC	SGK	17q21.2-q22	None	No	271 - 291
SGK237	2	14	Cat	Protein Kinas	CAMK	EMK	3p24.2-p21.3	527=S (ccctgagaccaS) ss727804; 82=K (ttgcttccttgK) ss1891216	AA430250.1, AI149647.1	None
SGK248	3	15	FL	Protein Kinas	CKI	CKI	15q21.3	None	BF308810.1, BE732523.1, BE900116.1	None
SGK223	4	16	FL	Protein Kinas	Other	AUR	8p22-p23	731=R (ccggcccacccccR) ss1379407; 2757=S (ccctgcccacacaS) ss1379408	BF820929.1, BF525647.1	3852 - 3874
SGK269	5	17	FL	Protein Kinas	Other	AUR	15q23	None	AW860071.1, AW862325.1, AA654785.1, AA679240	None
SGK258	6	18	FL	Protein kinas	Other	MLK	15	None	BF796030.1,BF333774.1, BF333812.1, AW502995.1	None
SGK382	7	19	FL	Protein Kinas	Other	Unique	4q24	3464=R (actattcacaatgcR) ss1252; 3336=W (acagatttagcattW) ss1512702	BE271267.1, BE518284.1, AA702160.1, AI808761.1	3540 - 3559
SGK424	8	20	FL	Protein kinas	Other	Unique	19q12-q13.3	None	No	None
SGK251	9	21	FL	Protein Kinas	TK	RTK-11	11	None	No	159 - 187
SGK307	10	22	FL	Protein kinas	TK	Unique	17q21.2-q22	2460=R (aagagagtttacaR) ss1928275	AL040739.1, AL040738.1, AI066732.1	None
SGK119	11	23	partial	PK-like	GCyc	GCyc	11q14	None	BF513152.1	None
SGK387	12	24	FL	PK-like	Inositol kinas	PI3K	16	None	BF086811.1; BE694351.1; AI050717.1	None

Table 3 lists the extent and the boundaries of the kinase catalytic domains, and other protein domains. The column headings are: "Gene Name", "ID#na", "ID#aa", "FL/Cat", "PK Profile_start", "PK Profile_end", "Protein Kinase_start", "Protein Kinase_end", "Profile", and "Additional Domains". The contents of the first 4 columns (i.e.,. "Gene Name", "ID#na", "ID#aa", and "FL/Cat") are as described above for Table 1. "Profile Start", "Profile End", "Kinase Start" and "Kinase End" refer to data obtained using a Hidden-Markov Model to define catalytic range boundaries. The profile has a length of 261 amino acids, corresponding to the complete protein kinase catalytic domain. Proteins in which the profile recognizes a full length catalytic domain have a "Profile Start" of 1 and a "Profile End" of 261. Genes which have a partial catalytic domain will have a "Profile Start" of greater than 1 (indicating that the beginning of the kinase domain is missing, and/or a "Profile End" of less than 261 (indicating that the C-terminal end of the kinase domain is missing). The boundaries of the catalytic domain within the overall protein are noted in the "Kinase Start" and "Kinase End" columns. "Profile" indicates whether the HMMR search was done with a complete ("Global") or Smith Waterman ("Partial") model, as described below. Starting from a multiple sequence alignment of kinase catalytic domains, two hidden Markov models were built. One of them allows for partial matches to the catalytic domain; this is a "local" HMM, similar to Smith-Waterman alignments in sequence matching. The other "complete" model allows matches only to the complete catalytic domain; this is a "global" HMM similar toNeedleman-Wunsch alignments in sequence matching. The Smith Waterman local model is more specific, allowing for fragmentary matches to the kinase catalytic domain whereas the global "complete" model is more sensitive, allowing for remote homologue identification. The "additional domains" column lists the names and positions of domains within the protein sequence in addition to the protein kinase domain. These domains were identified using PFAM (http://pfam.wustl.edu/hmmsearch.shtml) models, a large collection of multiple sequence alignments and hidden Markov models covering many common protein domains. Version 5.5 of Pfam (Sept 2000) contains alignments and models for 2478 protein families (http://pfam.wustl.edu/faq.shtml). The PFAM alignments were downloaded from http://pfam.wustl.edu/hmmsearch.shtml and the HMMr searches were run locally on a Timelogic computer (TimeLogic Corporation, Incline Village,

NV). The PFAM accession number, the length in amino acids and the number of proteins used to build the profile are listed below.

Table 3
Protein Kinase Domains, Other Domains

Gene Name	ID#na	ID#aa	FL/Cat	PK Profile start	PK Profile end	Protein Kinase start	Protein Kinase end	Profile	Additional Domains
SGK216	1	13	partial	15	261	4	224	Global Sugen	Protein kinase C terminal domain 225-276
SGK237	2	14	Cat	1	261	370	638	Global Sugen	Armadillo/beta-catenin-like repeats (2): 78-118 & 131-171
SGK248	3	15	FL	1	261	44	299	Global Sugen	None
SGK223	4	16	FL	34	134	857	894	Local Sugen	None
SGK269	5	17	FL	169	261	1577	1666	Local Sugen	None
SGK258	6	18	FL	1	261	1280	1563	Global Sugen	Leucine Rich Repeats (12): 276-301; 302-323; 328-351; 352-373; 380-403; 404-425; 450-472; 473-494; 497-518; 578-598; 599-622; 1222-1244
SGK382	7	19	FL	1	261	18	273	Global Sugen	TBC domain 463-673; Rhodanese-like domain 778-883
SGK424	8	20	FL	173	261	420	498	Local Sugen	None
SGK251	9	21	FL	1	261	631	930	Global Sugen	Ephrin receptor ligand binding domain 34-207; Fibronectin type III domain (2): 332-425 & 440-527; SAM domain (Sterile alpha motif) 859-1023
SGK307	10	22	FL	29	124	266	369	Local Sugen	Ankyrin (2) 55-87 & 88-120
SGK119	11	23	partial	PK-like	PK-like	PK-like	PK-like	NA	Adenylate and Guanylate cyclase catalytic domain 89-273
SGK387	12	24	FL	PK-like	PK-like	PK-like	PK-like	NA	Phosphatidylinositol 3- and 4-kinases 409-677; FATC domain 1878-1911

Table 4 describes the results of Smith Waterman similarity searches (Matrix: Pam100; gap open/extension penalties 12/2) of the amino acid sequences against the NCBI database of non-redundant protein sequences (http://www.ncbi.nlm.nih.gov/Entrez/protein.html). The column headings are: "Gene Name", "ID#na", "ID#aa", "FL/Cat", "Superfamily", "Group", "Family", "Pscore", "aa_length", "aa_ID_match", "%Identity", "%Similar", "ACC#_nraa_match", and "Description". The contents of the first 7 columns (i.e.,. "Gene Name", "ID#na", "ID#aa", "FL/Cat", "Superfamily", "Group", and "Family") are as described above for Table 1. "Pscore" refers to the Smith Waterman probability score. This number approximates the chance that the alignment occurred by chance. Thus, a very low number, such as 2.10E-64, indicates that there is a very significant match between the query and the database target. "aa_length" refers to the length of the protein in amino acids. "aa_ID_match" indicates the number of amino acids that were identical in the alignment. "% Identity" lists the percent of nucleotides that were identical over the aligned region. "% Similarity" lists the percent of amino acids that were similar over the alignment. "ACC#nraa_match" lists the accession number of the most similar protein in the NCBI database of non-redundant proteins. "Description" contains the name of the most similar protein in the NCBI database of non-redundant proteins.. In cases where the best hit is not very informative, a second Smith-Waterman hit is listed (the entries are right aligned for second entries).

Table 4

Smith Waterman

Gene Name	IDNa	IDAa	FL/Cat	Superfamily	Group	Family	Pscore	aa length	aa ID match	% Identity	% Similar	ACC# nraa match	Description
SGK216	1	13	no	Protein Kinase	AGC	YSK	########	281	271	96	99	NP 003162.1	Ribosomal protein S6 kinase, 70kD, polypeptide 1; serine/threonine kinase 14 alpha [Homo sapiens]
SGK237	2	14	Cat	Protein Kinase	CAMK	EMK	4.80E-71	854	227	40	61	P41851	STK D1044.3 (C. elegans)
SGK248	3	15	FL	Protein Kinase	CK1	CK1	########	422	422	100	100	BAB17639.1	Casein kinase 1, gamma 1 [Homo sapiens]
SGK223	4	16	FL	Protein Kinase	Other	AUR	3.40E-43	1244	159	50	65	BAB15008.1	(AK024783) unnamed protein product [Homo sapiens]
SGK269	5	17	FL	Protein Kinase	Other	AUR	########	1746	329	100	100	BAB15008.1	(AK024783) unnamed protein product [Homo sapiens]
SGK258	6	18	FL	Protein Kinase	Other	MLK	########	2014	253	100	100	BAB15547.1	(AK026772) unnamed protein product [Homo sapiens]
SGK258	6	18	FL	Protein Kinase	Other	MLK	1.10E-17	2014	133	29	49	AAF56793.1	(AE003731) CG5483 gene product [Drosophila melanogaster]
SGK382	7	19	FL	Protein Kinase	Other	Unique	########	893	483	85	95	AAF28980.1	HSPC302 [Homo sapiens]
SGK382	7	19	FL	Protein Kinase	Other	Unique	########	893	378	44	58	AAF45993.1	CG4041 gene product [Drosophila melanogaster]
SGK424	8	20	FL	Protein Kinase	Other	Unique	0.888764	560	23	35	60	NP 002812.1	PTK7 protein tyrosine kinase 7 [Homo sapiens]
GK251_EPHA	9	21	FL	Protein Kinase	TK	RTK-11	0	1035	1002	97	99	NP 031964.1	Eph receptor A8 [Mus musculus]
SGK307	10	22	FL	Protein Kinase	TK	Unique	9.20E-16	1428	64	27	48	P42588	TYROSINE-PROTEIN KINASE SRK1 [Spongilla lacustris]
SGK307	10	22	FL	Protein Kinase	TK	Unique	4.20E-14	1428	62	25	49	NP 002022.1	FRK (fyn-related kinase [Homo sapiens]
SGK119	11	23	no	PK-like	GCyc	GCyc	########	273	215	79	88	P51839	OLFACTORY GUANYLYL CYCLASE GC-D [Rattus norvegicus]
SGK387	12	24	FL	PK-like	inositol kinas	PI3K	0	1911	1302	100	100	BAA24851.1	(AB007881) KIAA0421 [Homo sapiens]
SGK387	12	24	FL	PK-like	inositol kinas	PI3K	9.00E-45	1911	189	28	48	NP 004948.1	FKBP binding protein 12-rapamycin associated protein 2 [Homo sapiens]

Table 5 gives results of a PCR screen of 48 human cDNA sources for 3 of the kinases represented in this application. A plus sign (+) indicates the presence of a band on an agarose gel of the expected size for the target kinase. A negative sign (-) indicates that the PCR product of the expected size was absent. The genes represented on this table are: SGK269 (SEQ ID NO: 5); SGK307 (SEQ ID NO: 10); and SGK387 (SEQ ID NO: 12). The columns in table 5 are as follows: "Tissue_name", "RNA_source" ("Clontech": from Clontech Inc; "Sugen": from in-house sources; "NCI": derived in-house from from human tumor cell lines), "Tumor_type" (tissue from which tumor is derived), "Species", and "Tumor_description".

Table 5
PCR Expression Analysis

Tissue name	RNA source	Tumor type	Species	Tumor description	SGK289 SEQID 8	SGK307 SEQ 10	SGK397 SEQID 12
fetal liver- h	Clontech		H		-	-	-
thymus,h	Clontech		H		+	-	-
pancreas - h	Clontech		H		+	-	-
pituitary gland - h	Clontech		H		-	-	-
placenta - h	Clontech		H		+	-	-
prostate, h	Clontech		H		+	+	-
salivary gl. - h	Clontech		H		+	-	-
skeletal muscle - h	Clontech		H		+	-	-
small intestine - h	Clontech		H		+	-	-
spinal cord - h	Clontech		H		+	-	-
Spleen - h	Clontech		H		+	-	-
stomach -h	Clontech		H		-	-	-
thyroid gland - h	Clontech		H		-	-	-
trachea - h	Clontech		H		+	-	-
uterus - h	Clontech		H		+	-	-
adrenal gland - h	Clontech		H		+	-	-
fetal brain - h	Clontech		H		+	-	-
fetal kidney - h	Clontech		H		+	-	-
fetal lung - h	Clontech		H		-	-	-
heart - h	Clontech		H		-	-	-
kidney - h	Clontech		H		-	-	-
liver - h	Clontech		H		-	-	-
lung - h	Clontech		H		-	-	-
lymph node - h	Clontech		H		-	+	-
Heart - h	Clontech		H		+	-	-
HPAEC	Sugen		H	Renal proximal tubule epithelial cells	+	-	-
RPTEC	Sugen		H	Mammary epithelial cells	-	-	-
HMEC	Sugen		H	Coronary artery endothelial cells	+	-	-
HCAEC	Sugen		H	Coronary artery endothelial cells	+	-	-
458 medullo RNA	Sugen		H	Neuroblastoma	+	-	-
A549/ATCC	NCI	LUNG	H	Lung carcinoma	+	-	-
MDA-MB-231	NCI	BREAST	H	Breat adenocarcinoma, pleural effusion	+	-	-
Hs 578T	NCI	BREAST	H	Ductal carcinoma	+	-	-
MCF-7/ADR-RES	NCI	BREAST	H	Breast adenocarcinoma	+	-	+
Malme-3M	NCI	MELANOMA	H	Malignant melanoma, metastatic to lung	+	-	-
A498	NCI	KIDNEY	H	Kidney carcinoma	-	-	-
COLO 205	NCI	COLON	H	Colon adenocarcinoma	+	-	+
CCRF-CEM	NCI	LEUKEMIA	H	ALL Acute lymphoblastic leukemia	+	+	+
SF-539	NCI	CNS	H	Glioblastoma	+	-	+
SF-295	NCI	CNS	H	Glioblastoma	+	-	+
U251	NCI	CNS	H	Glioblastoma	+	-	+
SNB-19	NCI	CNS	H	Glioblastoma	+	-	+
OVCAR-4	NCI	OVARY	H	Ovary adenocarcinoma	+	-	-
OVCAR-3	NCI	OVARY	H	Ovary adenocarcinoma	+	-	-
TCGP	GP	TESTIS	H	Testicular carcinoma	-	-	-
HMEC	GP		H	Coronary artery endothelial cells	-	-	-
HOP-62	NCI	LUNG	H	Lung adenocarcinoma	+	-	-
NCI-H522	NCI	LUNG	H	Lung adenocarcinoma	-	-	-

EXAMPLES

The examples below are not limiting and are merely representative of various aspects and features of the present invention. The examples below demonstrate the isolation and characterization of the nucleic acid molecules according to the invention, as well as the polypeptides they encode.

EXAMPLE 1: Identification and Characterization of Genomic Fragments Encoding Protein Kinases

Materials and Methods

Novel kinases were identified from the Celera human genomic sequence databases, and from the public Human Genome Sequencing project (http://www.ncbi.nlm.nih.gov/) using a hidden Markov model (HMMR) built with 70 mammalian and yeast kinase catalytic domain sequences. These sequences were chosen from a comprehensive collection of kinases such that no two sequences had more than 50% sequence identity. The genomic database entries were translated in six open reading frames and searched against the model using a Timelogic Decypher box with a Field programmable array (FPGA) accelerated version of HMMR2.1. The DNA sequences encoding the predicted protein sequences aligning to the HMMR profile were extracted from the original genomic database. The nucleic acid sequences were then clustered using the Pangea Clustering tool to eliminated repetitive entries. The putative protein kinase sequences were then sequentially run through a series of queries and filters to identify novel protein kinase sequences. Specifically, the HMMR identified sequences were searched using BLASTN and BLASTX against a nucleotide and amino acid repository containing known human protein kinases and all subsequent new protein kinase sequences as they are identified. The output was parsed into a spreadsheet to facilitate elimination of known genes by manual inspection. Two models were developed, a "complete" model and a "partial" or Smith Waterman model. The partial model was used to identify sub-catalytic kinase domains, whereas the complete model was used to identify complete catalytic domains. The selected hits were then queried using BLASTN against the public nrna and EST databases to confirm they are indeed unique. In some cases the novel genes were judged to be orthologues of previously identified rodent or vertebrate protein kinases.

Many of the sequences filed in the provisional patents did not contain the entire coding sequence. Extension of partial DNA sequences to encompass the full-length open-reading frame was carried out by several methods. Iterative blastn searching of the cDNA databases listed in Table 6 was used to find cDNAs that extended the genomic sequences. "LifeGold" databases are from Incyte Genomics, Inc (http://www.incyte.com/). NCBI databases are from the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/). All blastn searches were conducted using a blosum62 matrix, a penalty for a nucleotide mismatch of –3 and reward for a nucleotide match of 1. The gapped blast algorithm is described in: Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402).

Extension of partial DNA sequences to encompass the full-length open-reading frame was also carried out by iterative searches of genomic databases. The first method made use of the Smith-Waterman algorithm to carry out protein-protein searches of the closest homologue or orthologue to the partial. The target databases consisted of Genscan [Chris Burge and Sam Karlin "Prediction of Complete Gene Structures in Human Genomic DNA", JMB (1997) 268(1):78-94)] and open-reading frame (ORF) predictions of all human genomic sequence derived from the human genome project (HGP) as well as from Celera. The complete set of genomic databases searched is shown in Table 7 below. Genomic sequences encoding potential extensions were further assessed by blastp analysis against the NCBI nonredundant to confirm the novelty of the hit. The extending genomic sequences were incorporated into the cDNA sequence after removal of potential introns using the Seqman program from DNAStar. The default parameters used for Smith-Waterman searches were as shown next. Matrix: PAM100; gap-opening penalty: 12; gap extension penalty: 2. Genscan predictions were made using the Genscan program as detailed in Chris Burge and Sam Karlin "Prediction of Complete Gene Structures in Human Genomic DNA", JMB (1997) 268(1):78-94). ORF predictions from genomic DNA were made using a standard 6-frame translation.

Another method for defining DNA extensions from genomic sequence used iterative searches of genomic databases through the Genscan program to predict exon splicing [Burge and Karlin, JMB (1997) 268(1):78-94)]. These predicted genes were

then assessed to see if they represented "real" extensions of the partial genes based on homology to related kinases.

Another method involved using the Genewise program (http://www.sanger.ac.uk/Software/Wise2/) to predict potential ORFs based on homology to the closest orthologue/homologue. Genewise requires two inputs, the homologous protein, and genomic DNA containing the gene of interest. The genomic DNA was identified by blastn searches of Celera and Human Genome Project databases. The orthologs were identified by blastp searches of the NCBI non-redundant protein database (NRAA). Genewise compares the protein sequence to a genomic DNA sequence, allowing for introns and frameshifting errors.

Table 6: Databases used for cDNA-based sequence extensions

Database	Database Date
LifeGold templates	Jan 2001
LifeGold compseqs	Jan 2001
LifeGold compseqs	Jan 2001
LifeGold compseqs	Jan 2001
LifeGold fl	Jan 2001
LifeGold flft	Jan 2001
NCBI human Ests	Jan 2001
NCBI murine Ests	Jan 2001
NCBI nonredundant	Jan 2001

Table 7: Databases used for genomic-based sequence extensions

Database	Number of entries	Database Date
Celera v. 1-5	5,306,158	Jan 19/00
Celera v. 6-10	4,209,980	Mar24/00
Celera v. 11-14	7,222,425	Apr 24/00
Celera v. 15	243,044	May14/00
Celera v. 16-17	25,885	Apr 04/00
Celera Assembly 5 (release 25h)	479,986	Jan /01
HGP Phase 0	3,189	Nov 1/00
HGP Phase 1	20,447	Jan 1/01
HGP Phase 2	1,619	Jan 1/01
HGP Phase 3	9,224	Jan 1/01
HGP Chromosomal assemblies	2759	Aug 1/01

Results:

The sources for the sequence information used to extend the genes in the provisional patents are listed below. For genes that were extended using Genewise, the accession numbers of the protein ortholog and the genomic DNA are given. (Genewise uses the ortholog to assemble the coding sequence of the target gene from the genomic sequence). The amino acid sequences for the orthologs were obtained from the NCBI non-redundant database of proteins (http://www.ncbi.nlm.nih.gov/Entrez/protein.html). The genomic DNA came from two sources: Celera and NCBI-NRNA, as indicated below. cDNA sources are also listed below. All of the genomic sequences were used as input for Genscan predictions to predict splice sites [Burge and Karlin, JMB (1997) 268(1):78-94)]. Abbreviations: HGP: Human Genome Project; NCBI, National Center for Biotechnology Information.

SGK216 (SEQ ID NO: 1, encoding SEQ ID NO: 13)

Genewise orthologs: S12906.

Genomic DNA sources:Celera:160000117943196, 160000117455552; Human Genome Project: AC053481, gi|8099920

Notes: HGP phase 1 contig gi|8099920 is in 26 unordered pieces. Tblastn with human and rat S6 kinases allowed the contigs to be partially ordered, and genewise was run on that reordered contig, using the rat (gi|125696) and human (gi|4506737) genes. 12 AA polymorphisms were seen with respect to the human gene, of which two were stops. Comparison with Celera individual reads allowed correction of the sequence in the vicinity of the stops. Alternative Splice forms: The sequence KRNAASLGAGGPGDAGEV may be missing in one splice form, as shown by public ESTs gi|9868731 and gi|9868695.

SGK237 (SEQ ID NO: 2, encoding SEQ ID NO: 14)

Genewise orthologs: P41951 and NP_009410.

Genomic DNA sources:Celera:181000001058619; Human Genome Project:AC024936.

Notes: The original HMM hit was blasted against Celera_Asm5g, the original HMM hit aligned with genomic contig 181000001058619, 200 kb of which was used for genewise/genscan/sym4 predictions. Genewise was run with YLK3_CAEL as the model and the result extended the original sequence prediction. The genewise prediction was supported by Blastx vs. NCBI_nonredundant, confirming homology to YLK3_CAEL. NCBI EST AAA50735.1 extended the genewise prediction in the 5' direction while maintaining homology to YLK3_CAEL. No further extensions were possible by 5' EST walking. 92 kb of genomic sequence 5' to the original 200kb genomic sequence was run with a genewise prediction using YLK3_CAEL as the model. This genewise prediction extended the ORF in the 5' direction another 11 AA. No further 5' extension was predicted by genewise and no further EST hits were found. Genscan predicted 5 CDS's in this 92 kb stretch of genomic sequence. CDS 1 contains the gene for sodium bicarbonate cotransporter 3, defining a 5' boundary for SGK237. CDS 5 predicts several exons. When assembled, CDS 5 shares homology to YLK3_CAEL through AA 125 of YLK3_CAEL (AA 42 of SGK237). CDS 5 also overlaps with EST AAA50735.1 contiging it to the extended prediction. All EST sequences were corrected before use by blast against Celera_Asm5g and HGP_s databases._The Incyte EST 458972.1 extended the prediction in the 3' direction while maintaining homology to YLK3_CAEL. Genscan CDS3 predicts 9 exons, 5 of which

align to Incyte EST 458972.1, supporting the extension. Blastn of Incyte EST 458972.1 vs. EST databases fails to hit other ESTs. All EST sequences were corrected before use by blast against Celera_Asm5g and HGP_s databases.

SGK248 (SEQ ID NO: 3, encoding SEQ ID NO: 15)

Genewise orthologs: Q62761.

Genomic DNA sources:Celera:90000640246623; Human Genome Project:AC062023

This gene was cloned as a full length gene from a human testis cDNA library and fully sequenced. The gene was identified by BLASTN searching a database of arrayed genes. The sequence in this patent represents this physical reagent.

SGK223 (SEQ ID NO: 4, encoding SEQ ID NO: 16)

Genewise orthologs: AAF46188 and NP_015115.

Genomic DNA sources: Celera:301457013; Human Genome Project:AC068353, gi|10437181, gi|9082332 and gi|8516017.

The original HMM hit was blasted against HGP data (no long Celera assemblies exist for this), and the contigs gi|9082332 and gi|8516017 were used. Genewise was run with gi|10437181 and the result was extended iteratively using overlapping ESTs and long extensions of the open reading frame from genomic sequence. The ESTs used were Incyte 232675.1, 208779.1 and public ESTs gi|6701806, gi|12066443. EST sequences were corrected before use by Blast against HGP database. Most of protein is low-complexity, 20% identity to many other proteins, but comparison to EST databases shows mouse, rat, pig and cow ESTs of high similarity covering most of the gene.

SGK269 (SEQ ID NO: 5, encoding SEQ ID NO: 17)

Genewise orthologs: AAF46188 and NP_010182.

Genomic DNA sources:Celera:90000641936169; Human Genome Project:AC016693

cDNA sources: Incyte: 312916.1, 7394590H1, 048878.1; dbEST V87108

SGK258 (SEQ ID NO: 6, encoding SEQ ID NO: 18)

Genewise orthologs: T33475, AAF55793 and AAD18109.

Genomic DNA sources:Celera:181000000853046; Human Genome Project:AC019254, AC020578

cDNA Sources: Incyte 7483337CB1.

SGK382 (SEQ ID NO: 7, encoding SEQ ID NO: 19)

Genewise orthologs: AAF45995 and NP_013714.

Genomic DNA sources:Celera:92000003647415; Human Genome Project:AP001820

Notes: Genscan and genewise with homologs gi|7290543 and gi|7496925 was carried out contig 92000003647415. The genewise result was extended and corrected with the overlap cDNA gi|6841253 and ESTs, including gi|9342903, gi|9144919, gi|7113593. EST sequen corrected by comparison to the genomic sequence.

SGK424 (SEQ ID NO: 8, encoding SEQ ID NO: 20)

Genewise orthologs: Q91048 and AAD42856.

Genomic DNA sources:Celera:92000004018261; Human Genome Project:AC002398

The final sequence is derived from Genescan prediction using genomic sequence referenced above.

SGK251 (SEQ ID NO: 9, encoding SEQ ID NO: 21)

Genewise orthologs: NP_031964.

Genomic DNA sources:Celera:90000642861512; Human Genome Project:U90093

This gene covers several genomic contigs. The mouse ortholog EPHA6 gi|6679661 gene was used to find genomic matching fragments, which were found in contigs: 90000642861512, 181000001006202, 181000001006202, AC023837.12 and AC021156.4 Genewise was run using the mouse homolog, gi|6679661 and separate genewise fragments were pieced together based on their homology to mouse. Blast against EST databases produced a 3' UTR from gi|5395573, whose sequence was corrected by comparing with genomic, and an N-terminal extension was generated from incyte EST 1135879.1, again, corrected with genomic sequence. Celera read 39000025421751 continued the sequence upstream, keeping both an open reading frame and homology with the mouse cDNA.

SGK307 (SEQ ID NO: 10, encoding SEQ ID NO: 22)

Genewise orthologs: Q91987.

Genomic DNA sources:Celera:181000001871106; Human Genome Project:AC011195

This gene was cloned as a full length gene from a human testis library. The clone was identified by BLASTN of a database representing arrayed plasmids from a testis cDNA library. The sequence in this patent represents the sequence of this physical reagent.

SGK119 (SEQ ID NO: 11, encoding SEQ ID NO: 23)

Genewise orthologs: P51839.

Genomic DNA sources:Celera:90000628251764; Human Genome Project:AP001189

The original HMM hit was blasted against Celera_Asm5g, the original HMM hit aligned with genomic contig 90000628251764, 200 kb of which was used for genewise/genscan/sym4 predictions. Genewise was run with CYGX_r as the model and the result extended the original sequence prediction to 831 AA. The genewise prediction was supported by Blastx vs. NCBI_nonredundant, confirming homology to CYGX_r. Genscan predictions failed to further extend SGK249. Genewise and Genscan predictions were blastn vs. all EST and cDNA databases. Several hits were found (LGtemplatesJAN2001: AAC42057 / LGcompseqJAN2001: 761235T6 and 761235H1 / Celera ORF: hCT1641294 / NCBI_Human_EST: gi:11598331).

SGK387 (SEQ ID NO: 12, encoding SEQ ID NO: 24)

Genewise orthologs: AAF46207 and AAD48773.

Genomic DNA sources:Celera:90000641304227; Human Genome Project:AC025289 , AC026472

The Genewise prediction was extended approximately 7 kb 3' using AB007881.1, and app 150 bp 5' with the public EST BF371317; EST sequence were corrected by comparison to sequence.

SGK216, SEQ ID NO: 1 (encoding SEQ ID NO: 13) is 843 nucleotides long. The open reading frame starts at position 1 and ends at position 843, giving an ORF length of 843 nucleotides. The predicted protein is 281 amino acids long. This sequence is a partial kinase catalytic domain. It is classified as (superfamily/group/family): Protein kinase, AGC, S6K. This gene maps to chromosomal position 17q21.2-q22. Amplification of this chromosomal position has been assosciated with the following human diseases: ovarian cancer (at position17q21-qter, with a frequency of 3/47); liver and lymph node metastases (at position17q21-q22, with a frequency of 3/14);

and breast carcinoma (at position17q22-q25, with a frequency of 8/101). (Knuutila, et al.). This gene does not contain mapped candidate single nucleotide polymorphisms. No ESTs representing this gene in were not found in dbEST. This gene has repetitive sequence at nucleotide positions 271 - 291 (sequence: gacctgaagccggagaatatc).

SGK237, SEQ ID NO: 2 (encoding SEQ ID NO: 14) is 2562 nucleotides long. The open reading frame starts at position 1 and ends at position 2562, giving an ORF length of 2562 nucleotides. The predicted protein is 854 amino acids long. This sequence contains a complete kinase catalytic domain. It is classified as (superfamily/group/family): protein kinase, CAMK, EMK. This gene maps to chromosomal position 3p24.2-p21.3. This chromosomal position has not been associated with human diseases. This gene contains two candidate single nucleotide polymorphisms at position 527 (527=S, tgaggaccctgagaccaS, dbSNP ss727804); and position 82 (82=K, ttccttgcttccttgK, dbSNP ss1891216). The SNP at position 527 changes the amino acid sequence. When nucleotide 527 is a G, then amino acid 176 is a serine. When nucleotide 527 is a C, then amino acid 176 is a threonine. The SNP at position 82 also changes the amino acid sequence. When nucleotide 82 is a G, then amino acid 28 is a glycine. When nucleotide 82 is a T, then amino acid 28 is a cysteine. These changes in the protein sequence could affect its activity within the cell. ESTs for this gene in the public domain (dbEST) are: AA430250.1 and AI149647.1. This gene does not contain repetitive sequences as defined above (over 21 bp with 100% match in multiple genes).

SGK248, SEQ ID NO: 3 (encoding SEQ ID NO: 15) is 1269 nucleotides long. The open reading frame starts at position 1 and ends at position 1266, giving an ORF length of 1266 nucleotides. The predicted protein is 422 amino acids long. This sequence is full length (start methionine to stop codon). It is classified as (superfamily/group/family): protein kinase, CKI, CKI. This gene maps to chromosomal position 15q21.3. This chromosomal position has not been associated with human diseases. This gene does not contain mapped candidate single nucleotide polymorphisms. ESTs for this gene in the public domain (dbEST) are: BF308810.1, BE732523.1, BE900116.1. This gene does not contain repetitive sequences as defined above (over 21 bp with 100% match in multiple genes). There are several C-terminal splice variants.

SGK223, SEQ ID NO: 4 (encoding SEQ ID NO: 16) is 4169 nucleotides long. The open reading frame starts at position 1 and ends at position 3732, giving an ORF length of 3732 nucleotides. The predicted protein is 1244 amino acids long. This sequence is full length (start methionine to stop codon). It is classified as (superfamily/group/family): protein kinase, other, AUR. This gene maps to chromosomal position 8p22-p23. Amplification of this chromosomal position has been assosciated with stomach carcinoma (at position8p22-p23, with a frequency of 3/58). (Knuutila, et al.). This gene contains two candidate single nucleotide polymorphisms: at position 731 (731=R, gccccggcccacccccR, dbSNP ss1379407; and position 2757 (2757=S, gcctgccctgcccacacaS, dbSNP ss1379406). The SNP at 731 changes the amino acid sequence. When nucleotide 731 is an A, then amino acid 244 is a glutamine. When nucleotide 731 is a G, then amino acid 244 is an arginine. This is a non-conservative change, since the arginine is charged and the glutamine is neutral. The SNP at position 2757 also changes the amino acid sequence. When nucleotide 2757 is a C, then amino acid 919 is a histidine. When nucleotide 2757 is a G, then amino acid 919 is a glutamine. This change in amino acid sequence could alter the proteins function within the cell. ESTs for this gene in the public domain (dbEST) are: BF820929.1, BF525647.1. This gene has repetitive sequence at the following nucleotide positions: 3852 - 3874, (sequence: taaaatatataaatatatatata).

SGK269, SEQ ID NO: 5 (encoding SEQ ID NO: 17) is 5241 nucleotides long. The open reading frame starts at position 1 and ends at position 5238, giving an ORF length of 5238 nucleotides. The predicted protein is 1746 amino acids long. This sequence is full length (start methionine to stop codon). It is classified as (superfamily/group/family): protein kinase, other, AUR. This gene maps to chromosomal position 15q23. Amplification of this chromosomal position has been assosciated with osteosarcoma (at position15q23-qter, with a frequency of 1/31). (Knuutila, et al.). This gene does not contain mapped candidate single nucleotide polymorphisms. ESTs for this gene in the public domain (dbEST) are: AW860071.1, AW862325.1, AA654785.1, and AA679240. This gene does not contain repetitive sequences as defined above (over 21 bp with 100% match in multiple genes).

SGK258, SEQ ID NO: 6 (encoding SEQ ID NO: 18) is 6042 nucleotides long. The open reading frame starts at position 1 and ends at position 6042, giving an ORF length of 6042 nucleotides. The predicted protein is 2014 amino acids long. This sequence is full length (start methionine to stop codon). It is classified as (superfamily/group/family): Protein kinase, Other, MLK. This gene maps to chromosome 15. This chromosomal position is too imprecise to assign specific disease associations. This gene does not contain mapped candidate single nucleotide polymorphisms. ESTs for this gene in the public domain (dbEST) are: BF796030.1,BF333774.1, BF333812.1, and AW502995.1. This gene does not contain repetitive sequences as defined above (over 21 bp with 100% match in multiple genes).

SGK382, SEQ ID NO: 7 (encoding SEQ ID NO: 19) is 3915 nucleotides long. The open reading frame starts at position 358 and ends at position 3036, giving an ORF length of 2679 nucleotides. The predicted protein is 893 amino acids long. This sequence is full length (start methionine to stop codon). It is classified as (superfamily/group/family): Protein kinase, Other, Unique. This gene maps to chromosomal position 4q24. This chromosomal position has been associated with testicular cancer (at position14q12-qter, with a frequency of 2/11). (Knuutila, et al.). This gene contains two candidate single nucleotide polymorphisms at the following postions: at position 3484 (3484=R, aactgagaaactattcacaatgcR, dbSNP ss1252; and position 3336 (3336=W, acaagacagatttagcattW, dbSNP ss1512702). Both SNPs are in the 3' untranslated region of the gene. ESTs for this gene in the public domain (dbEST) are: BE271267.1, BE513284.1, AA702160.1, AI808761.1. This gene has repetitive sequence at the following nucleotide positions: 3540 - 3559, (sequence gtgtgtgtgtgtgtgtgtat).

SGK424, SEQ ID NO: 8 (encoding SEQ ID NO: 20) is 1683 nucleotides long. The open reading frame starts at position 1 and ends at position 1680, giving an ORF length of 1680 nucleotides. The predicted protein is 560 amino acids long. This sequence is full length (start methionine to stop codon). It is classified as (superfamily/group/family): Protein kinase, Other, Unique. This gene maps to chromosomal position 19q12-q13.3 Amplification of this chromosomal position has been assosciated with small cell lung cancer (at position19q13.1, with a frequency of

10/35). (Knuutila, et al.). This gene does not contain mapped candidate single nucleotide polymorphisms. . ESTs for this gene in the public domain (dbEST) were not found. . This gene does not contain repetitive sequences as defined above (over 21 bp with 100% match in multiple genes).

SGK251, SEQ ID NO: 9 (encoding SEQ ID NO: 21) is 3903 nucleotides long. The open reading frame starts at position 280 and ends at position 3387, giving an ORF length of 3108 nucleotides. The predicted protein is 1036 amino acids long. This sequence is full length (start methionine to stop codon). It is classified as (superfamily/group/family): Protein kinase, TK, RTK-11. This gene maps to chromosome 11. This chromosomal position is too imprecise to assign specific disease associations. This gene does not contain mapped candidate single nucleotide polymorphisms. No ESTs for this gene were found in dbEST. This gene has repetitive sequence at the following nucleotide positions: 159 – 187; (sequence: ggaggaggaagaggaggaggaggaagaag).

SGK307, SEQ ID NO: 10 (encoding SEQ ID NO: 22) is 4739 nucleotides long. The open reading frame starts at position 123 and ends at position 4406, giving an ORF length of 4284 nucleotides. The predicted protein is 1428 amino acids long. This sequence is full length (start methionine to stop codon). It is classified as (superfamily/group/family): Protein kinase, TK, Unique. This gene maps to chromosomal position 17q21.2-q22. Amplification of this chromosomal position has been assosciated with ovarian cancer (at position17q21-qter, with a frequency of 3/47). (Knuutila, et al.). This gene contains a candidate single nucleotide polymorphism at position 2460 (2460=R, acttcaagagagtttacaR, dbSNP ss1928275). This SNP alters the amino acid sequence of the protein. When nucleotide 2460 is an A, them amino acid 780 is an asparagine. When nucleotide 2460 is a G, then amino acid 780 is an aspartic acid. This non-conservative change in amino acid composition could alter the function of the encoded protein. ESTs for this gene in the public domain (dbEST) are: AL040739.1, AL040738.1, AI066732.1. This gene does not contain repetitive sequences as defined above (over 21 bp with 100% match in multiple genes).

SGK119, SEQ ID NO: 11 (encoding SEQ ID NO: 23) is 819 nucleotides long. The open reading frame starts at position 1 and ends at position 819, giving an ORF length of 819 nucleotides. The predicted protein is 273 amino acids long. This sequence is a partial kinase catalytic domain. It is classified as (superfamily/group/family): PK-like, GCyc, GCyc. This gene maps to chromosomal position 11q14. Amplifications involving this chromosomal position has been associated with bladder carcinoma (at position11q14-q22, with a frequency of 1/16). (Knuutila, et al.). This gene does not contain mapped candidate single nucleotide polymorphisms. One EST for this gene in the public domain (dbEST) was identified: BF513152.1. This gene does not contain repetitive sequences as defined above (over 21 bp with 100% match in multiple genes).

SGK387, SEQ ID NO: 12 (encoding SEQ ID NO: 24) is 5736 nucleotides long. The open reading frame starts at position 1 and ends at position 5733, giving an ORF length of 5733 nucleotides. The predicted protein is 1911 amino acids long. This sequence is full length (start methionine to stop codon). It is classified as (superfamily/group/family): PK-like, Inositol kinase, PI3K. This gene maps to chromosome 16. This chromosomal position is too imprecise to assign specific disease associations. This gene does not contain mapped candidate single nucleotide polymorphisms. ESTs for this gene in the public domain (dbEST) are: BF086811.1; BE694351.1; AI050717.1. This gene does not contain repetitive sequences as defined above (over 21 bp with 100% match in multiple genes).

EXAMPLE 2: Expression Analysis of Polypeptides of the Invention

Expression Analysis

The gene expression patterns for selected genes were studied using techniques a PCR screen of 48 human tissues (this technique does not yield quantitative expression levels between tissues, but does identify which tissues express the gene at a level detectable by PCR and which do not).

PCR Screening

Screening for expression sources by PCR from ds cDNA templates: PCR screening of cDNAs from various sources allows identification of tissues that express the gene of interest. We screened 48 different human cDNA sources for gene

expression. The genes were: SGK269 (SEQ ID NO: 5); SGK307 (SEQ ID NO: 10); and SGK387 (SEQ ID NO: 12). The 48 tissues and cell lines, listed in column one of Table 5, were as follows: fetal liver, thymus, pancreas, pituitary gland, placenta, prostate, salivary gland, skeletal muscle, small intestine, spinal cord , Spleen , stomach -h, thyroid gland , trachea , uterus , adrenal gland , fetal brain , fetal kidney , fetal lung , heart , kidney , liver , lung , lymph node , Heart , HPAEC, RPTEC, HMEC, HCAEC, 458 medullo RNA, A549/ATCC , MDA-MB-231, Hs 578T, MCF-7/ADR-RES, Malme-3M, A498, COLO 205, CCRF-CEM, SF-539, SF-295, U251, SNB-19, OVCAR-4, OVCAR-3, TCGP, HMEC, HOP-62, NCI-H522.

Preparation of dscDNA templates

dscDNA templates were prepared by PCR amplification of symmetrically-tagged reverse transcriptase sscDNA products generated as described in detail under Materials and Methods for the Tissue Array Gene Expression protocol. The tissue sources amplified are listed in Table 5. The amplification conditions were as follows: per 200 microl of PCR reaction, added 100 microl of Premix TaKaRa ExTaq, 20.0 microl of pwo DNA polymerase (1/10 dilution made as follows: 1 microl pwo (5 units/microl), 1 microl 10x PCR buffer with 20 mM MgSO4, 8 microl water), 4.0 microl sscDNA template (reverse transcriptase product), 8.0 microl 10 pmoles/microl (10 microM) primer (AAGCAGTGGTAACAACGCAGAGT) (1.0 microM final conc.) and 68.0 microl H_20. The reaction was amplified according to the following regiment: hot start (95^0C for 1 min), 95^0C for 1 min, 24 cycles, 95^0C for 20 s, 65^0C for 30 s, 68^0C for 6 min, 68^0C for 10 min, 1 cycle and 4^0C forever. Following the PCR reaction, 5-10 microl of product were applied to an agarose gel together with 1kb ladder size standards to assess the yield and uniformity of the product. A positive sign (+) in the table indicates the presence of the PCR product at the expected size. Products were cut out for sequence verification.

The oligonucleotides used to screen the cDNA sources, and the size of the PCR product, are listed below.

SGK269 5' CAGGTGTGTCTGCTGCTCTTACAGC
3' CGGTGCTGCAGAATTTTCACAATAC
size: 715bp

SGK307 5' CAGCAGCGGTCCCAGTTCCCAG
3' GTAAGCAGAAATAAACTCCCAACAAC
size: 570bp

SGK387 5' GCGCCTGCTGCGGTTGCTCGTGAAG
3' CACCAGAGGGTCGTACACAAAGGCC
size: 1.71kb

Results

SGK269 (SEQ ID NO: 5) was successfully identified by PCR from the following human tissues/cell lines: thymus, pancreas, placenta, prostate, salivary gland, skeletal muscle, small intestine, spinal cord , spleen, stomach, trachea, uterus, adrenal gland, fetal brain , fetal kidney, Heart, HPAEC, HMEC, HCAEC, 458 medullo RNA, A549/ATCC , MDA-MB-231, Hs 578T, MCF-7/ADR-RES, Malme-3M, COLO 205, CCRF-CEM, SF-539, SF-295, U251, SNB-19, OVCAR-4, OVCAR-3, HOP-62. This gene is widely expressed in both tumor and normal tissue.

SGK307 (SEQ ID NO: 10), was successfully identified by PCR from the following human tissues/cell lines: salivary gland, lymph node, CCRF-CEM leukemia cell line. This pattern suggests that the gene may have a role in hematopoiesis or immune functions.

SGK387 (SEQ ID NO: 12) was successfully identified by PCR from the following human tumor-derived cell lines: MCF-7/ADR-RES, COLO 205, CCRF-CEM, SF-539, SF-295, U251, and SNB-19. The expression was restricted to tumor derived cell lines, suggesting that this gene may play a role in human cancer.

EXAMPLE 3: Chromosomal Localization of Protein Kinases

Materials and Methods

Several sources were used to find information about the chromosomal localization of each of the genes described in this patent. First, .cytogenetic map locations of these contigs were found in the title or text of their Genbank record, or by inspection through the NCBI human genome map viewer (http://www.ncbi.nlm.nih.gov/cgi-bin/Entrez/hum_srch?). Alternatively, the accession number of a genomic contig (identified by BLAST against NRNA) was

used to query the Entrez Genome Browser (http://www.ncbi.nlm.nih.gov/PMGifs/Genomes/MapViewerHelp.html), and the cytogenetic localization was read from the NCBI data. A thorough search of available literature for the cytogenetic region is also made using Medline (http://www.ncbi.nlm.nih.gov/PubMed/medline.html). References for association of the mapped sites with chromosomal amplifications found in human cancer can be found in: Knuutila, et al., Am J Pathol, 1998, 152:1107-1123.

Results

The chromosomal regions for mapped genes are listed Table 2, and are discussed in the section Nucleic Acids above. The chromosomal positions were cross-checked with the Online Mendelian Inheritance in Man database (OMIM, **http://www.ncbi.nlm.nih.gov/htbin-post/Omim**)., which tracks genetic information for many human diseases, including cancer. References for association of the mapped sites with chromosomal abnormalities found in human cancer can be found in: Knuutila, et al., Am J Pathol, 1998, 152:1107-1123. A third source of information on mapped positions was searching published literature (at NCBI, http://www.ncbi.nlm.nih.gov/entrez/query.fcgi) for documented association of the mapped position with human disease.

EXAMPLE 4: Candidate Single Nucleotide Polymorphisms (SNPs)
Materials and Methods

The most common variations in human DNA are single nucleotide polymorphisms (SNPs), which occur approximately once every 100 to 300 bases. Because SNPs are expected to facilitate large-scale association genetics studies, there has recently been great interest in SNP discovery and detection. Candidate SNPs for the genes in this patent were identified by blastn searching the nucleic acid sequences against the public database of sequences containing documented SNPs (dbSNP, at NCBI, http://www.ncbi.nlm.nih.gov/SNP/snpblastpretty.html). dbSNP accession numbers for the SNP-containing sequences are given. SNPs were also identified by comparing several databases of expressed genes (dbEST, NRNA) and genomic sequence (i.e., NRNA) for single basepair mismatches. The results are shown in Table 1, in the column labeled "SNPs". These are candidate SNPs – their actual

frequency in the human population was not determined. The code below is standard for representing DNA sequence:

G = Guanosine
A = Adenosine
T = Thymidine
C = Cytidine
R = G or A, puRine
Y = C or T, pYrimidine
K = G or T, Keto
W = A or T, Weak (2 H-bonds)
S = C or G, Strong (3 H-bonds)
M = A or C, aMino
B = C, G or T (i.e., not A)
D = A, G or T (i.e., not C)
H = A, C or T (i.e., not G)
V = A, C or G (i.e., not T)
N = A, C, G or T, aNy
X = A, C, G or T

complementary	G A T C R Y W S K M B V D H N X
DNA	+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
strands	C T A G Y R S W M K V B H D N X

For example, if two versions of a gene exist, one with a "C" at a given position, and a second one with a "T: at the same position, then that position is represented as a Y, which means C or T. SNPs may be important in identifying heritable traits associated with a gene.

Results

The results of SNP identification are reviewed in the Nucleic Acids section above.

EXAMPLE 5: Predicted Proteins

SGK216 (SEQ ID NO: 1) encodes SEQ ID NO: 13, a protein that is 281 amino acids long. It is classified as (superfamily/group/family): Protein Kinase, AGC, S6K. The kinase domain in this protein matches the hidden Markov profile for a full length kinase domain of 261 amino acids from profile position 15 to profile position 261. The position of the kinase catalytic region within the encoded protein is from amino acid 4 to amino acid 224. The results of a Smith Waterman search of the public database of amino acid sequences (NRAA) with this protein sequence yielded the following results: Pscore = 1.40E-105; number of identical amino acids = 271; percent identity = 96%; percent similarity = 99%; the accession number of the most similar entry in NRAA is NP_003152.1; the name or description, and species, of the most similar protein in NRAA is: Ribosomal protein S6 kinase, 70kD, polypeptide 1; serine/threonine kinase 14 alpha [Homo sapiens]. Domains other than the kinase catalytic domain identified within this protein are: protein kinase C terminal domain, at amino acid position 225-278

SGK237 (SEQ ID NO: 2) encodes SEQ ID NO: 14, a protein that is 854 amino acids long. It is classified as (superfamily/group/family): Protein Kinase, CAMK, EMK. The kinase domain in this protein matches the hidden Markov profile for a full length kinase domain of 261 amino acids from profile position 1 to profile position 261. The position of the kinase catalytic region within the encoded protein is from amino acid 370 to amino acid 636. The results of a Smith Waterman search of the public database of amino acid sequences (NRAA) with this protein sequence yielded the following results: Pscore = 4.80E-71; number of identical amino acids = 227; percent identity = 40%; percent similarity = 61%; the accession number of the most similar entry in NRAA is P41951; the name or description, and species, of the most similar protein in NRAA is: STK D1044.3 [C. elegans]. Domains other than the kinase catalytic domain identified within this protein are: Armadillo/beta-catenin-like repeats (2): amino acids 78-118 & 131-171

SGK248 (SEQ ID NO: 3) encodes SEQ ID NO: 15, a protein that is 422 amino acids long. It is classified as (superfamily/group/family): Protein Kinase, CKI, CKI. The kinase domain in this protein matches the hidden Markov profile for a full length kinase domain of 261 amino acids from profile position 1 to profile position

261. The position of the kinase catalytic region within the encoded protein is from amino acid 44 to amino acid 299. The results of a Smith Waterman search of the public database of amino acid sequences (NRAA) with this protein sequence yielded the following results: Pscore = 6.80E-195; number of identical amino acids = 422; percent identity = 100%; percent similarity = 100%; the accession number of the most similar entry in NRAA is BAB17839.1; the name or description, and species, of the most similar protein in NRAA is: Casein kinase 1, gamma 1 [Homo sapiens].

SGK223 (SEQ ID NO: 4) encodes SEQ ID NO: 16, a protein that is 1244 amino acids long. It is classified as (superfamily/group/family): Protein Kinase, Other, AUR. The kinase domain in this protein matches the hidden Markov profile for a full length kinase domain of 261 amino acids from profile position 34 to profile position 134. The position of the kinase catalytic region within the encoded protein is from amino acid 857 to amino acid 994. The results of a Smith Waterman search of the public database of amino acid sequences (NRAA) with this protein sequence yielded the following results: Pscore = 3.40E-43; number of identical amino acids = 159; percent identity = 50%; percent similarity = 65%; the accession number of the most similar entry in NRAA is BAB15006.1; the name or description, and species, of the most similar protein in NRAA is: (AK024793) unnamed protein product [Homo sapiens].

SGK269 (SEQ ID NO: 5) encodes SEQ ID NO: 17, a protein that is 1746 amino acids long. It is classified as (superfamily/group/family): Protein Kinase, Other, AUR. The kinase domain in this protein matches the hidden Markov profile for a full length kinase domain of 261 amino acids from profile position 169 to profile position 261. The position of the kinase catalytic region within the encoded protein is from amino acid 1577 to amino acid 1666. The results of a Smith Waterman search of the public database of amino acid sequences (NRAA) with this protein sequence yielded the following results: Pscore = 1.60E-166; number of identical amino acids = 329; percent identity = 100%; percent similarity = 100%; the accession number of the most similar entry in NRAA is BAB15006.1; the name or description, and species, of the most similar protein in NRAA is: (AK024793) unnamed protein product [Homo sapiens].

SGK258 (SEQ ID NO: 6) encodes SEQ ID NO: 18, a protein that is 2014 amino acids long. It is classified as (superfamily/group/family): Protein kinase, Other, MLK. The kinase domain in this protein matches the hidden Markov profile for a full length kinase domain of 261 amino acids from profile position 1 to profile position 261. The position of the kinase catalytic region within the encoded protein is from amino acid 1280 to amino acid 1563. The results of a Smith Waterman search of the public database of amino acid sequences (NRAA) with this protein sequence yielded the following results: Pscore = 2.50E-104; number of identical amino acids = 253; percent identity = 100%; percent similarity = 100%; the accession number of the most similar entry in NRAA is BAB15547.1; the name or description, and species, of the most similar protein in NRAA is: (AK026772) unnamed protein product [Homo sapiens]. Domains other than the kinase catalytic domain identified within this protein are: Leucine Rich Repeats (12): 278-301; 302-323; 329-351; 352-373; 380-403; 404-425; 450-472; 473-494; 497-519; 576-598; 599-622; 1222-1244

SGK382 (SEQ ID NO: 7) encodes SEQ ID NO: 19, a protein that is 893 amino acids lon classified as (superfamily/group/family): Protein Kinase, Other, Unique. The kinase domain in t protein matches the hidden Markov profile for a full length kinase domain of 261 amino acids fr profile position 1 to profile position 261. The position of the kinase catalytic region within the encoded protein is from amino acid 18 to amino acid 273. The results of a Smith Waterman sear the public database of amino acid sequences (NRAA) with this protein sequence yielded the foll results: Pscore = 1.80E-200; number of identical amino acids = 483; percent identity = 95%; perc similarity = 95%; the accession number of the most similar entry in NRAA is AAF28980.1; the or description, and species, of the most similar protein in NRAA is: HSPC302 [Homo sapiens]. Domains other than the kinase catalytic domain identified within this protein are: TBC domain, a acids 463-673; Rhodanese-like domain, amino acids 776-883. The TBC domain is found in rab-l GAPs (GYP6_YEAST and GYP7_YEAST, generally small proteins, though one worm protein contains this and two kinase domains). The rhodanese domain is found in rhodanese and in som phosphatases and ubiquitin C-terminal hydrolases.

SGK424 (SEQ ID NO: 8) encodes SEQ ID NO: 20, a protein that is 560 amino acids long. It is classified as (superfamily/group/family): Protein kinase, Other, Unique. The kinase domain in this protein matches the hidden Markov profile for a full length kinase domain of 261 amino acids from profile position 173 to profile

position 261. The position of the kinase catalytic region within the encoded protein is from amino acid 420 to amino acid 496. The results of a Smith Waterman search of the public database of amino acid sequences (NRAA) with this protein sequence yielded the following results: Pscore = 0.988764; number of identical amino acids = 23; percent identity = 35%; percent similarity = 60%; the accession number of the most similar entry in NRAA is NP_002812.1; the name or description, and species, of the most similar protein in NRAA is: PTK7 protein tyrosine kinase 7 [Homo sapiens].

SGK251 (SEQ ID NO: 9) encodes SEQ ID NO: 21, a protein that is 1036 amino acids long. It is classified as (superfamily/group/family): Protein Kinase, TK, RTK-11. The kinase domain in this protein matches the hidden Markov profile for a full length kinase domain of 261 amino acids from profile position 1 to profile position 261. The position of the kinase catalytic region within the encoded protein is from amino acid 631 to amino acid 930. The results of a Smith Waterman search of the public database of amino acid sequences (NRAA) with this protein sequence yielded the following results: Pscore = 0; number of identical amino acids = 1002; percent identity = 97%; percent similarity = 99%; the accession number of the most similar entry in NRAA is NP_031964.1; the name or description, and species, of the most similar protein in NRAA is: Eph receptor A6 [Mus musculus]. Domains other than the kinase catalytic domain identified within this protein are: Ephrin receptor ligand binding domain, amino acids 34-207; Fibronectin type III domain (2): amino acids 332-425 & 440-527; SAM domain (Sterile alpha motif), amino acids 959-1023.

SGK307 (SEQ ID NO: 10) encodes SEQ ID NO: 22, a protein that is 1428 amino acids long. It is classified as (superfamily/group/family): Protein kinase, TK, Unique. The kinase domain in this protein matches the hidden Markov profile for a full length kinase domain of 261 amino acids from profile position 29 to profile position 124. The position of the kinase catalytic region within the encoded protein is from amino acid 266 to amino acid 369. The results of a Smith Waterman search of the public database of amino acid sequences (NRAA) with this protein sequence yielded the following results: Pscore = 0.000161; number of identical amino acids = 49; percent identity = 31%; percent similarity = 52%; the accession number of the most similar entry in NRAA is NP_057062.1; the name or description, and species, of the most similar protein in NRAA is: Putative protein-tyrosine kinase [Homo

sapiens]. Domains other than the kinase catalytic domain identified within this protein are: Ankyrin (2), amino acids 55-87 & 88-120.

SGK119 (SEQ ID NO: 11) encodes SEQ ID NO: 23, a protein that is 273 amino acids long. It is classified as (superfamily/group/family): PK-like, GCyc, GCyc. This protein is related to the protein kinase family but has a substantially different catalytic region. The results of a Smith Waterman search of the public database of amino acid sequences (NRAA) with this protein sequence yielded the following results: Pscore = 2.30E-126; number of identical amino acids = 215; percent identity = 79%; percent similarity = 88%; the accession number of the most similar entry in NRAA is P51839; the name or description, and species, of the most similar protein in NRAA is: OLFACTORY GUANYLYL CYCLASE GC-D [Rattus norviegicus]. The guanyalte cyclase catalytic domain identified in this protein by PFAM searching is located between amino acids 89 and 273.

SGK387 (SEQ ID NO: 12) encodes SEQ ID NO: 24, a protein that is 1911 amino acids long. It is classified as (superfamily/group/family): PK-like, Inositol kinase, PI3K. This protein is related to the protein kinase family but has a substantially different catalytic region. The results of a Smith Waterman search of the public database of amino acid sequences (NRAA) with this protein sequence yielded the following results: Pscore = 0; number of identical amino acids = 1302; percent identity = 100%; percent similarity = 100%; the accession number of the most similar entry in NRAA is BAA24851.1; the name or description, and species, of the most similar protein in NRAA is: (AB007881) KIAA0421 [Homo sapiens]. The phosphatidylinositol 3- and 4-kinase domain identified by PFAM searching is located between amino acids 408-677. The FATC domain identified by PFAM searching is located between amino acids 1879 and 1911.

EXAMPLE 6: Classification of polypeptides exhibiting kinase activity among defined groups

AGC Group

Potential biological and clinical implications of the novel AGC group protein kinases are described next.

The partial SGK216 (SEQ ID NO: 13) belonging to the S6K family of AGC group kinases is 96% identical over a 281 amino acid region to human ribosomal protein S6 kinase (NP_003152.1). The family of human ribosomal S6 protein kinases consists of at least 8 members (RSK1, RSK2, RSK3, RSK4, MSK1, MSK2, p70S6K and p70S6Kb). Ribosomal protein S6 protein kinases play important pleotropic functions, among them is a key role in the regulation of mRNA translation during protein biosynthesis (*Eur J Biochem* 2000 Nov; 267(21):6321-30, Exp Cell Res. 1999 Nov 25;253 (1):100-9, *Mol Cell Endocrinol* 1999 May 25;151(1-2):65-77). The phosphorylation of the S6 ribosomal protein by p70S6 has also been implicated in the regulation of cell motility (*Immunol Cell Biol* 2000 Aug;78(4):447-51) and cell growth (*Prog Nucleic Acid Res Mol Biol* 2000;65:101-27), and hence, may be important in tumor metastasis, the immune response and tissue repair. SGK216 may represent an additional member of the family of S6 kinases with a potential role in cancer, inflammation, as well as other disease conditions.

CAMK Group

Potential biological and clinical implications of the novel CAMK group protein kinases are described next. The Partial SGK237 (SEQ ID NO: 14) belonging to the EMK family of CAMK kinases is 40% identical over a 854 amino acid region to C. elegans STK YLK3 (P41951). SGK237 and C. elegans STK YLK3 have considerable homology over their extracatalytic regions and, hence, may represent orthologous proteins that share related biological functions. The function of the C. elegans STK YLK3 protein is unknown. The next closest match to SGK237 is the human NIMA (never in mitosis gene a)-related kinase 6 (NRK6) (NP_055212.1) with an amino acid sequence identity of 35% over 261 amino acids. The homology between SGK237 and NRK6 is confined to the catalytic region. The functions of NRK6 (9q33. 3-->q34.11) is unknown. NRK6, together with NEK1, NEK2 and NEK3 are proteins that share a high degree of catalytic domain sequence homology to each other, and that have a low, but significant catalytic domain homology to the NIMA kinase from the filamentous fungus Aspergillus nidulans. The pivotal role of NIMA in the regulation of mitosis in Aspergillus is well documented *Prog Cell Cycle Res* 1995;1:187-205 *Cytogenet Cell Genet* 1999;87(3-4):271-2).

The homology between SGK237 and *c. elegans* STK YLK3 extends beyond the catalytic domain boundaries of these proteins. Particularly significant is the

presence of two armadillo/beta-catenin-like repeats in the human protein and one such repeat in the worm protein located immediately N-terminal to the catalytic domain. The armadillo/beta-catenin-like repeat, initially characterized in the drosophila wingless protein, is found in a number of diverse signaling proteins. This repeat is responsible for protein-protein interactions that take place at intercellular contacts and within the cellular cytoskeleton (*Int Rev Cytol* 1999;186:179-224). The armadillo/beta-catenin-like oligodimerization repeat plays a key role in the function of the adenomatous polyposis coli (APC) tumor suppressor protein (*Biochim Biophys Acta* 1997 Jun 7;1332(3):F127-47). The C. elegans STK YLK3 protein also features 12 EGF-like repeats or EB module (EBM) domains located C-terminal to the catalytic domain not found in human SGK237. The EBM domain whose function is unknown has only been predicted to exist only in proteins encoded by invertebrate genomes. SGK237 may play a role in intra- and inter-cellular signaling important for cell growth, differentiation and/or apoptosis.

Casein Kinase (CK1) group

Potential biological and clinical implications of the novel CK1 group protein kinases are described next.

The full-length SGK248 (SEQ ID NO: 15) belonging to the EMK family of CAMK group kinases is 100% to human casein kinase 1, gamma 1 (NP_071331.1). Alternative splicing at the C-terminus for SGK248 may exist based on the identification a cDNA clone (SGK248C_variant) that diverges in the 3' region relative to SGK248. The Cterminus of SGK248 is SLRTVTAEHYDVNNSAIWHRGRGT (encoded by the sequence (AGCCTTAGGACTGTTACAGCTGAGCATTATGATGTTAACAACTCAGCCAT CTGGCACAGGGGAAGAGGCACC). The Cterminus of the SGK248C_variant is EIFVD (encoded by the sequence GAAATTTTTGTGGAT). Alternative splicing at the Cterminus may result in altered protein function.

"Other" group

Potential biological and clinical implications of the novel protein kinases belonging to the Other group are described next.

The full-length SGK223 (SEQ ID NO: 16) and SGK269 (SEQ ID NO: 17) belong to the Aurora family of (Other) group protein kinases. SGK223 displays 50%

amino acid sequence identity over 1244 amino acids to a hypothetical human protein (BAB15006.1). SGK269 displays 100% amino acid sequence identity over a partial region (329 amino acids) of its entire length (1746 amino acids) to the same human hypothetical protein (BAB15006.1). SGK223 and SGK269 have low but significant homology over their catalytic domain to a pair of drosophila hypothetical proteins (CG4523 and CG10967) of unknown funtion. The classification of SGK223 and SGK269 into the Aurora family of kinases stems from a short stretch of catalytic homology to Ip11p from saccharomyces cerevisiae (NP_015115.1). Members of the Aurora/Ipl1p family of serine/threonine kinases include Aurora1 (NM_004217, Aurora2 (AF008551) and Aurora3 (AF059681) and have a demonstrated role as mitotic regulators (*Trends Cell Biol* 1999 Nov;9(11):454-9).

The full-length SGK258 (SEQ ID NO: 18) belongs to the mixed-lineage kinase (MLK) family of (Other) group protein kinases. SGK258 displays100% amino acid sequence identity over a partial region (253 amino acids) of its entire length (2014 amino acids) to a hypothetical human protein (BAB15547.1). The next closest matches to SGK258 are the drosophila melanogaster CG5483 gene product (AAF55793.1) (29% over 2014 amino acids). The function of the Drosophila melanogaster CG5483 gene product is unknown. SGK258 features a 12 leucine-rich repeats that may be responsible for protein-protein interactions.

The full-length SGK382 (SEQ ID NO: 19) and SGK424 (SEQ ID NO: 20) belong to the Unique family of (Other) group protein kinases. Unique family members have low sequence homology over their catalytic region relative to other kinases found within the Other group . SGK382 (SEQ ID NO: 19) displays 44% amino acid sequence identity over 893 amino acids to the drosophila CG4041 gene product (AAF45995.1). The function of this invertebrate protein is unknown. SGK424 displays 35% amino acid sequence identity over 560 amino acids to human PTK7 protein tyrosine kinase 7 (NP_002812.1), a transmembrane glycoprotein receptor tyrosine kinase expressed in human colon carcinoma cells (*Oncogene* 1995 Nov 16;11(10):2179-84). The short length of the match between SGK424 and human PTK7 (23 amino acids) precludes drawing any conclusions about the potential function of SGK424 based on this sequence comparison.

TK Group

The full-length SGK251 (SEQ ID NO: 21) is a new tyrosine kinase related to the ephrin family of receptors. SGK251 displays 97% amino acid sequence identity over 1036 amino acids to the murine Eph receptor A6 (NP_031964.1), and hence, is likely to be the human orthologue of the rodent kinase. SGK251 features the canonical motifs associated with the ephrin receptor proteins, which, in addition to the kinase domain and a potential transmembrane domain include: a ligand-binding domain, two fibronectin type III domains and a SAM (sterile alpha motif) domain. The ephrin receptors (Eph) and their ligands (ephrins) function in axon guidance, neuron-target interactions and other synaptic functions of the nervous system.

The full-length SGK307 (SEQ ID NO: 22) is a new tyrosine kinase belonging to the unique family of tyrosine kinases. Unique family members have low sequence homology over their catalytic region relative to other kinases found within the tyrosine kinase group. SGK307 displays 28% amino acid sequence identity over 1428 amino acids to a hypothetical protein from arabidopsis thaliana, protein kinase homolog F24O1.13. The next closest homologs to SGK307 exhibit similar homologies to the top hit and includ primarily tyrosine kinases from diverse organisms such as src-related tyrosine kinase from Spongilla lacustris (AAB22579.1), human fyn-related kinase (NP_002022.1) and Xenopus laevis c-Src kinase (AAC05835.1). A Hidden-Markov profile analysis on the 1428 amino acid full-length SGK307 protein predicts two ankyrin domains at positions 55-87 and 88-120 and a kinase domain at position 295-493. The role of ankyrin repeats protein-protein interactions is well documented (*J Cell Sci* 2000 Jun;113 (Pt 11):1851-6). Analysis for potential membrane-spanning regions in SGK307 (http://psort.nibb.ac.jp/cgi-bin/runpsort.pl) predicts a potential nuclear or Golgi apparatus transmembrane domain. A hydrophobic region (position 59-68) consistent with this prediction is observed using the protean DNAStar program. SGK307 may function as a tyrosine kinase associated with intracellular organelles such as the Golgi apparatus and/or the nucleus. Through its ankyrin repeat SGK307 could conceivably be responsible for interacting with and phosphorylating Golgi proteins such as tankyrase, an ankyrin repeat-containing protein demonstrated to be a MAPK substrate which is associated with GLUT4 vesicles and is important in insulin secretion (*J Biol Chem* 2000 Dec 8;275(49):38437-44).

EXAMPLE 7: Classification of polypeptides exhibiting kinase-like activity among defined groups

Guanylate Cyclase Group

The Partial SGK119 (SEQ ID NO: 23) belonging to the guanylate-cyclase family of guanylate-cyclase group kinases is 79% identical over a 273 amino acid region to rat olfactory guanylyl cyclase gc-d (P51839). A hidden Markov profile analysis on the the partial 273 amino acid SGK119 amino acid sequence predicted partial kinase catalytic domain (position 1-19) followed by a complete adenylate /guanylate cyclase catalytic domain at positions 89-273. The presence of the bipartitate catalytic structure is a characteristic feature of a number of adenylate /guanylate cyclases. SGK119 may represent an additional member of the odorant receptor guanyl cyclases expressed in olfactory sensory neurons. The apparently very restricted pattern of expression associated with SGK119 judged from the lack of cDNA clone representation in diverse cDNAlibraries argues in favor of this possibility. Guanylate cyclase function is not restricted to olfaction. Being responsible for generating cGMP, guanylate cyclases play a ubiquitous role as upstream regulators of cGMP-dependant signaling cascades involving cGMP-dependent protein kinases, cGMP-regulated phosphodiesterases, and cyclic nucleotide-gated ion channels. Increasing evidence points towards an involvment of guanylate cyclases in pathophysiological processes that involve vascular smooth muscle motility, intestinal fluid and electrolyte homeostasis, and retinal phototransduction (*Pharmacol Rev* 2000 Sep;52(3):375-414). As a new member of the growing family of the adenylate /guanylate cyclases, SGK119 may play a role in these pathophysiological events. In addition, exogenous and endogenous nitric oxide stimulates on myocardial functions such as contraction, relaxation and heart rate through a signaling cascade that involves guanylate cyclases (*Cardiovasc Res* 1999 Mar;41(3):514-23). SGK119 may, consequently, play a role in cardiovascular disease.

Inositol kinase group

The full-length SGK387 (SEQ ID NO: 24) belonging to the PI3K (phosphatidylinositol kinase) family of inositol kinase group kinases is 100% identical over a partial region (1302 amino acids) of its entire length (1911 amino acids) to the human protein KIAA0421 (BAA24851.1). SGK387 is also identical over its entire

length to the partial-length protein, lambda/iota protein kinase C-interacting protein (U32581). The next closest matches to SGK387include various proteins from diverse species that are likely to be close orthologue or close homologs given the length of their alignment to SGK387. These include

The CG4549 gene product from drosophila (AAF46207.1), nonsense-mediated mRNA decay protein SMG-1 (AAD48773.1) from Caenorhabditis elegans, TOR-like protein (AAG43422.1) from arabidopsis thaliana, mutant drr1-1 protein (AAB66881.1) from Saccharomyces cerevisiae and human FK506 binding protein 12-rapamycin associated protein 1 (FRAP) (NP_004949.1).

A hidden Markov model analysis of the full-length 1911 amino acid SGK387 revealed a potential phosphatidylinositol 3- and 4-kinase catalytic domain at position 408-677 and towards the Cterminus a FATC domain (FRAP, ATM, TRRAP Cterminal) at positions 1879-1911. The domain structure predicted for SGK387 is consistent with that observed in the human FRAP protein which features a central, long FATC domain, and at the C-terminus a phosphatidylinositol kinase domain that is immediately followed by a short FATC domain. A similar structure is seen in other members of this class of PI3 kinases such as the Tel1 and Mec1 from Saccharomyces cerevisiae and the human ataxia telangiectasia mutated (ATM) protein. Yeast Tel1 and Mec1, and human ATM are involved in telomere length regulation and cellular responses to DNA damage (*Proc Natl Acad Sci U S A* 2000 Dec 5;97(25):13749-54).

The role of ATM in familial and sporadic cancer has been well-documented (*Recent Results Cancer Res* 1998;154:156-73). Deletions in the chromosomal locus 1p36.2 harboring the ATM-related FRAP gene are found in human neuroblastomas (*Hum Genet* 1997 Apr;99(4):547-9). Like Tel1 from yeast and human ATM and FRAP, SGK387 is likely to play a role as a cell-cycle checkpoint protein. Mutations arising from SGK387 or disruptions in its signaling mechanisms may be associated with human neoplasia. In support of a potential cancer-related role for SGK387 is RT-PCR evidence presented in this application suggestive of higher expression levels of this gene in tumor versus normal cellular sources (table 5).

EXAMPLE 8: Isolation of cDNAs Encoding Mammalian Protein Kinases

Materials and Methods

Identification of novel clones

Total RNAs are isolated using the Guanidine Salts/Phenol extraction protocol of Chomczynski and Sacchi (P. Chomczynski and N. Sacchi, *Anal. Biochem.* 162, 156 (1987)) from primary human tumors, normal and tumor cell lines, normal human tissues, and sorted human hematopoietic cells. These RNAs are used to generate single-stranded cDNA using the Superscript Preamplification System (GIBCO BRL, Gaithersburg, MD; Gerard, GF *et al.* (1989), *FOCUS* 11, 66) under conditions recommended by the manufacturer. A typical reaction uses 10 μg total RNA with 1.5 μg oligo$(dT)_{12-18}$ in a reaction volume of 60 μL. The product is treated with RNaseH and diluted to 100 μL with H_20. For subsequent PCR amplification, 1-4 μL of this sscDNA is used in each reaction.

Degenerate oligonucleotides are synthesized on an Applied Biosystems 3948 DNA synthesizer using established phosphoramidite chemistry, precipitated with ethanol and used unpurified for PCR. These primers are derived from the sense and antisense strands of conserved motifs within the catalytic domain of several protein kinases. Degenerate nucleotide residue designations are: N = A, C, G, or T; R = A or G; Y = C or T; H = A, C or T not G; D = A, G or T not C; S = C or G; and W = A or T.

PCR reactions are performed using degenerate primers applied to multiple single-stranded cDNAs. The primers are added at a final concentration of 5 μM each to a mixture containing 10 mM TrisHCl, pH 8.3, 50 mM KCl, 1.5 mM $MgCl_2$, 200 μM each deoxynucleoside triphosphate, 0.001% gelatin, 1.5 U AmpliTaq DNA Polymerase (Perkin-Elmer/Cetus), and 1-4 μL cDNA. Following 3 min denaturation at 95 °C, the cycling conditions are 94 °C for 30 s, 50 °C for 1 min, and 72 °C for 1 min 45 s for 35 cycles. PCR fragments migrating between 300-350 bp are isolated from 2% agarose gels using the GeneClean Kit (Bio101), and T-A cloned into the pCRII vector (Invitrogen Corp. U.S.A.) according to the manufacturer's protocol.

Colonies are selected for mini plasmid DNA-preparations using Qiagen columns and the plasmid DNA is sequenced using a cycle sequencing dye-terminator kit with AmpliTaq DNA Polymerase, FS (ABI, Foster City, CA). Sequencing reaction products are run on an ABI Prism 377 DNA Sequencer, and analyzed using the BLAST alignment algorithm (Altschul, S.F. *et al.*, *J.Mol.Biol.* 215: 403-10).

Additional PCR strategies are employed to connect various PCR fragments or ESTs using exact or near exact oligonucleotide primers. PCR conditions are as described above except the annealing temperatures are calculated for each oligo pair using the formula: Tm = 4(G+C)+2(A+T).

Isolation of cDNA clones:

Human cDNA libraries are probed with PCR or EST fragments corresponding to kinase-related genes. Probes are ^{32}P-labeled by random priming and used at $2x10^6$ cpm/mL following standard techniques for library screening. Pre-hybridization (3 h) and hybridization (overnight) are conducted at 42 oC in 5X SSC, 5X Denhart's solution, 2.5% dextran sulfate, 50 mM $Na_2PO_4/NaHPO_4$, pH 7.0, 50% formamide with 100 mg/mL denatured salmon sperm DNA. Stringent washes are performed at 65 °C in 0.1X SSC and 0.1% SDS. DNA sequencing was carried out on both strands using a cycle sequencing dye-terminator kit with AmpliTaq DNA Polymerase, FS (ABI, Foster City, CA). Sequencing reaction products are run on an ABI Prism 377 DNA Sequencer.

EXAMPLE 9: Expression Analysis of Mammalian Protein Kinases

Materials and Methods

Northern blot analysis

Northern blots are prepared by running 10 μg total RNA isolated from 60 human tumor cell lines (such as HOP-92, EKVX, NCI-H23, NCI-H226, NCI-H322M, NCI-H460, NCI-H522, A549, HOP-62, OVCAR-3, OVCAR-4, OVCAR-5, OVCAR-8, IGROV1, SK-OV-3, SNB-19, SNB-75, U251, SF-268, SF-295, SF-539, CCRF-CEM, K-562, MOLT-4, HL-60, RPMI 8226, SR, DU-145, PC-3, HT-29, HCC-2998, HCT-116, SW620, Colo 205, HTC15, KM-12, UO-31, SN12C, A498, CaKi1, RXF-393, ACHN, 786-0, TK-10, LOX IMVI, Malme-3M, SK-MEL-2, SK-MEL-5, SK-MEL-28, UACC-62, UACC-257, M14, MCF-7, MCF-7/ADR RES, Hs578T, MDA-MB-231, MDA-MB-435, MDA-N, BT-549, T47D), from human adult tissues (such as thymus, lung, duodenum, colon, testis, brain, cerebellum, cortex, salivary gland, liver, pancreas, kidney, spleen, stomach, uterus, prostate, skeletal muscle, placenta, mammary gland, bladder, lymph node, adipose tissue), and 2 human fetal normal

tissues (fetal liver, fetal brain), on a denaturing formaldehyde 1.2% agarose gel and transferring to nylon membranes.

Filters are hybridized with random primed [α^{32}P]dCTP-labeled probes synthesized from the inserts of several of the kinase genes. Hybridization is performed at 42 °C overnight in 6X SSC, 0.1% SDS, 1X Denhardt's solution, 100 μg/mL denatured herring sperm DNA with 1-2 x 10^6 cpm/mL of ^{32}P-labeled DNA probes. The filters are washed in 0.1X SSC/0.1% SDS, 65 °C, and exposed on a Molecular Dynamics phosphorimager.

Quantitative PCR analysis

RNA is isolated from a variety of normal human tissues and cell lines. Single stranded cDNA is synthesized from 10 μg of each RNA as described above using the Superscript Preamplification System (GibcoBRL). These single strand templates are then used in a 25 cycle PCR reaction with primers specific to each clone. Reaction products are electrophoresed on 2% agarose gels, stained with ethidium bromide and photographed on a UV light box. The relative intensity of the STK-specific bands were estimated for each sample.

DNA Array Based Expression Analysis

Plasmid DNA array blots are prepared by loading 0.5 μg denatured plasmid for each kinase on a nylon membrane. The [γ^{32}P]dCTP labeled single stranded DNA probes are synthesized from the total RNA isolated from several human immune tissue sources or tumor cells (such as thymus, dendrocytes, mast cells, monocytes, B cells (primary, Jurkat, RPMI8226, SR), T cells (CD8/CD4+, TH1, TH2, CEM, MOLT4), K562 (megakaryocytes). Hybridization is performed at 42 °C for 16 hours in 6X SSC, 0.1% SDS, 1X Denhardt's solution, 100 μg/mL denatured herring sperm DNA with 10^6 cpm/mL of [γ^{32}P]dCTP labeled single stranded probe. The filters are washed in 0.1X SSC/0.1% SDS, 65 °C, and exposed for quantitative analysis on a Molecular Dynamics phosphorimager.

EXAMPLE 10: Protein Kinase Gene Expression

Vector Construction

Materials and Methods

Expression Vector Construction

Expression constructs are generated for some of the human cDNAs including: a) full-length clones in a pCDNA expression vector; b) a GST-fusion construct containing the catalytic domain of the novel kinase fused to the C-terminal end of a GST expression cassette; and c) a full-length clone containing a Lys to Ala (K to A) mutation at the predicted ATP binding site within the kinase domain, inserted in the pCDNA vector.

The "K to A" mutants of the kinase might function as dominant negative constructs, and will be used to elucidate the function of these novel STKs.

EXAMPLE 11: Generation of Specific Immunoreagents to Protein Kinases

Materials and Methods

Specific immunoreagents are raised in rabbits against KLH- or MAP-conjugated synthetic peptides corresponding to isolated kinase polypeptides. C-terminal peptides were conjugated to KLH with glutaraldehyde, leaving a free C-terminus. Internal peptides were MAP-conjugated with a blocked N-terminus. Additional immunoreagents can also be generated by immunizing rabbits with the bacterially expressed GST-fusion proteins containing the cytoplasmic domains of each novel PTK or STK.

The various immune sera are first tested for reactivity and selectivity to recombinant protein, prior to testing for endogenous sources.

Western blots

Proteins in SDS PAGE are transferred to immobilon membrane. The washing buffer is PBST (standard phosphate-buffered saline pH 7.4 + 0.1% Triton X-100). Blocking and antibody incubation buffer is PBST +5% milk. Antibody dilutions varied from 1:1000 to 1:2000.

EXAMPLE 12: Recombinant Expression and Biological Assays for Protein Kinases

Materials and Methods

Transient Expression of Kinases in Mammalian Cells

The pcDNA expression plasmids (10 μg DNA/100 mm plate) containing the kinase constructs are introduced into 293 cells with lipofectamine (Gibco BRL). After 72 hours, the cells are harvested in 0.5 mL solubilization buffer (20 mM HEPES, pH 7.35, 150 mM NaCl, 10% glycerol, 1% Triton X-100, 1.5 mM $MgCl_2$, 1 mM EGTA, 2 mM phenylmethylsulfonyl fluoride, 1 μg/mL aprotinin). Sample aliquots are resolved by SDS polyacrylamide gel electrophoresis (PAGE) on 6% acrylamide/0.5% bis-acrylamide gels and electrophoretically transferred to nitrocellulose. Non-specific binding is blocked by preincubating blots in Blotto (phosphate buffered saline containing 5% w/v non-fat dried milk and 0.2% v/v nonidet P-40 (Sigma)), and recombinant protein was detected using the various anti-peptide or anti-GST-fusion specific antisera.

In Vitro Kinase Assays

Three days after transfection with the kinase expression constructs, a 10 cm plate of 293 cells is washed with PBS and solubilized on ice with 2 mL PBSTDS containing phosphatase inhibitors (10 mM $NaHPO_4$, pH 7.25, 150 mM NaCl, 1% Triton X-100, 0.5% deoxycholate, 0.1% SDS, 0.2% sodium azide, 1 mM NaF, 1 mM EGTA, 4 mM sodium orthovanadate, 1% aprotinin, 5 μg/mL leupeptin). Cell debris was removed by centrifugation (12000 x g, 15 min, 4 °C) and the lysate was precleared by two successive incubations with 50 μL of a 1:1 slurry of protein A sepharose for 1 hour each. One-half mL of the cleared supernatant was reacted with 10 μL of protein A purified kinase-specific antisera (generated from the GST fusion protein or antipeptide antisera) plus 50 μL of a 1:1 slurry of protein A-sepharose for 2 hr at 4 °C. The beads were then washed 2 times in PBSTDS, and 2 times in HNTG (20 mM HEPES, pH 7.5/150 mM NaCl, 0,1% Triton X-100, 10% glycerol).

The immunopurified kinases on sepharose beads are resuspended in 20 μL HNTG plus 30 mM $MgCl_2$, 10 mM $MnCl_2$, and 20 μCi [α^{32}P]ATP (3000 Ci/mmol). The kinase reactions are run for 30 min at room temperature, and stopped by addition of HNTG supplemented with 50 mM EDTA. The samples are washed 6 times in

HNTG, boiled 5 min in SDS sample buffer and analyzed by 6% SDS-PAGE followed by autoradiography. Phosphoamino acid analysis is performed by standard 2D methods on ^{32}P-labeled bands excised from the SDS-PAGE gel.

Similar assays are performed on bacterially expressed GST-fusion constructs of the kinases.

EXAMPLE 13: Demonstration Of Gene Amplification By Southern Blotting

Materials and Methods

Nylon membranes are purchased from Boehringer Mannheim. Denaturing solution contains 0.4 M NaOH and 0.6 M NaCl. Neutralization solution contains 0.5 M Tris-HCL, pH 7.5 and 1.5 M NaCl. Hybridization solution contains 50% formamide, 6X SSPE, 2.5X Denhardt's solution, 0.2 mg/mL denatured salmon DNA, 0.1 mg/mL yeast tRNA, and 0.2 % sodium dodecyl sulfate. Restriction enzymes are purchased from Boehringer Mannheim. Radiolabeled probes are prepared using the Prime-it II kit by Stratagene. The beta actin DNA fragment used for a probe template is purchased from Clontech.

Genomic DNA is isolated from a variety of tumor cell lines (such as MCF-7, MDA-MB-231, Calu-6, A549, HCT-15, HT-29, Colo 205, LS-180, DLD-1, HCT-116, PC3, CAPAN-2, MIA-PaCa-2, PANC-1, AsPc-1, BxPC-3, OVCAR-3, SKOV3, SW 626 and PA-1, and from two normal cell lines.

A 10 μg aliquot of each genomic DNA sample is digested with EcoR I restriction enzyme and a separate 10 μg sample is digested with Hind III restriction enzyme. The restriction-digested DNA samples are loaded onto a 0.7% agarose gel and, following electrophoretic separation, the DNA is capillary-transferred to a nylon membrane by standard methods (Sambrook, J. *et al* (1989) Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory).

EXAMPLE 14: Detection Of Protein-Protein Interaction Through Phage Display

Materials And Methods

Phage display provides a method for isolating molecular interactions based on affinity for a desired bait. cDNA fragments cloned as fusions to phage coat proteins are displayed on the surface of the phage. Phage(s) interacting with a bait are

enriched by affinity purification and the insert DNA from individual clones is analyzed.

T7 Phage Display Libraries

All libraries were constructed in the T7Select1-1b vector (Novagen) according to the manufacturer's directions.

Bait Presentation

Protein domains to be used as baits are generated as C-terminal fusions to GST and expressed in *E. coli*. Peptides are chemically synthesized and biotinylated at the N-terminus using a long chain spacer biotin reagent.

Selection

Aliquots of refreshed libraries (10^{10}-10^{12} pfu) supplemented with PanMix and a cocktail of *E. coli* inhibitors (Sigma P-8465) are incubated for 1-2 hrs at room temperature with the immobilized baits. Unbound phage is extensively washed (at least 4 times) with wash buffer.

After 3-4 rounds of selection, bound phage is eluted in 100 µL of 1% SDS and plated on agarose plates to obtain single plaques.

Identification of insert DNAs

Individual plaques are picked into 25 µL of 10 mM EDTA and the phage is disrupted by heating at 70 °C for 10 min. 2 µL of the disrupted phage are added to 50 µL PCR reaction mix. The insert DNA is amplified by 35 rounds of thermal cycling (94 °C, 50 sec; 50 °C, 1min; 72 °C, 1min).

Composition of Buffer

10x PanMix

5% Triton X-100

10% non-fat dry milk (Carnation)

10 mM EGTA

250 mM NaF

250 μg/mL Heparin (sigma)

250 μg/mL sheared, boiled salmon sperm DNA (sigma)

0.05% Na azide

Prepared in PBS

Wash Buffer

PBS supplemented with:

0.5% NP-40

25 μl g/mL heparin

PCR reaction mix

1.0 mL 10x PCR buffer (Perkin-Elmer, with 15 mM Mg)

0.2 mL each dNTPs (10 mM stock)

0.1 mL T7UP primer (15 pmol/μL) GGAGCTGTCGTATTCCAGTC

0.1 mL T7DN primer (15 pmol/μL) AACCCCTCAAGACCCGTTTAG

0.2 mL 25 mM $MgCl_2$ or $MgSO_4$ to compensate for EDTA

Q.S. to 10 mL with distilled water

Add 1 unit of Taq polymerase per 50 μL reaction

LIBRARY: T7 Select1-H441

EXAMPLE 15: FLK-1

An ELISA assay was conducted to measure the kinase activity of the FLK-1 receptor and more specifically, the inhibition or activation of TK activity on the FLK-1 receptor. Specifically, the following assay was conducted to measure kinase activity of the FLK-1 receptor in cells genetically engineered to express Flk-1.

Materials and Reagents

The following reagents and supplies were used:

1. Corning 96-well ELISA plates (Corning Catalog No. 25805-96);

2. Cappel goat anti-rabbit IgG (catalog no. 55641);

3. PBS (Gibco Catalog No. 450-1300EB);

4. TBSW Buffer (50 mM Tris (pH 7.2), 150 mM NaCl and 0.1% Tween-20);

5. Ethanolamine stock (10% ethanolamine (pH 7.0), stored at 4 °C);

6. HNTG buffer (20 mM HEPES buffer (pH 7.5), 150 mM NaCl, 0.2% Triton X-100, and 10% glycerol);

7. EDTA (0.5 M (pH 7.0) as a 100X stock);

8. Sodium orthovanadate (0.5 M as a 100X stock);

9. Sodium pyrophosphate (0.2 M as a 100X stock);

10. NUNC 96 well V bottom polypropylene plates (Applied Scientific Catalog No. AS-72092);

11. NIH3T3 C7#3 Cells (FLK-1 expressing cells);

12. DMEM with 1X high glucose L-Glutamine (catalog No. 11965-050);

13. FBS, Gibco (catalog no. 16000-028);

14. L-glutamine, Gibco (catalog no. 25030-016);

15. VEGF, PeproTech, Inc. (catalog no. 100-20) (kept as 1 μg/100 μl stock in Milli-Q dH_2O and stored at -20 °C);

16. Affinity purified anti-FLK-1 antiserum;

17. UB40 monoclonal antibody specific for phosphotyrosine (see, Fendley, *et al.*, 1990, *Cancer Research* 50:1550-1558);

18. EIA grade Goat anti-mouse IgG-POD (BioRad catalog no. 172-1011);

19. 2,2-azino-bis(3-ethylbenz-thiazoline-6-sulfonic acid (ABTS) solution (100 mM citric acid (anhydrous), 250 mM Na_2HPO_4 (pH 4.0), 0.5 mg/ml ABTS (Sigma catalog no. A-1888)), solution should be stored in dark at 4 °C until ready for use;

20. H_2O_2 (30% solution) (Fisher catalog no. H325);

21. ABTS/H_2O_2 (15 ml ABTS solution, 2 μl H_2O_2) prepared 5 minutes before use and left at room temperature;

22. 0.2 M HCl stock in H_2O;

23. dimethylsulfoxide (100%) (Sigma Catalog No. D-8418); and

24. Trypsin-EDTA (Gibco BRL Catalog No. 25200-049).

Protocol

The following protocol was used for conducting the assay:

1. Coat Corning 96-well ELISA plates with 1.0 μg per well Cappel Anti-rabbit IgG antibody in 0.1 M Na_2CO_3 pH 9.6. Bring final volume to 150 μl per well. Coat plates overnight at 4 °C. Plates can be kept up to two weeks when stored at 4 °C.

2. Grow cells in Growth media (DMEM, supplemented with 2.0 mM L-Glutamine, 10% FBS) in suitable culture dishes until confluent at 37 °C, 5% CO_2.

3. Harvest cells by trypsinization and seed in Corning 25850 polystyrene 96-well round bottom cell plates, 25.000 cells/well in 200 μl of growth media.

4. Grow cells at least one day at 37 °C, 5% CO_2.

5. Wash cells with D-PBS 1X.

6. Add 200 μl/well of starvation media (DMEM, 2.0 mM l-Glutamine, 0.1% FBS). Incubate overnight at 37 °C, 5% CO_2.

7. Dilute Compounds 1:20 in polypropylene 96 well plates using starvation media. Dilute dimethylsulfoxide 1:20 for use in control wells.

8. Remove starvation media from 96 well cell culture plates and add 162 μl of fresh starvation media to each well.

9. Add 18 μl of 1:20 diluted Compound dilution (from step 7) to each well plus the 1:20 dimethylsulfoxide dilution to the control wells (± VEGF), for a final dilution of 1:200 after cell stimulation. Final dimethylsulfoxide is 0.5%. Incubate the plate at 37 °C, 5% CO_2 for two hours.

10. Remove unbound antibody from ELISA plates by inverting plate to remove liquid. Wash 3 times with TBSW + 0.5% ethanolamine, pH 7.0. Pat the plate on a paper towel to remove excess liquid and bubbles.

11. Block plates with TBSW + 0.5% Ethanolamine, pH 7.0, 150 μl per well. Incubate plate thirty minutes while shaking on a microtiter plate shaker.

12. Wash plate 3 times as described in step 10.

13. Add 0.5 μg/well affinity purified anti-FLU-1 polyclonal rabbit antiserum. Bring final volume to 150 μl/well with TBSW + 0.5% ethanolamine pH 7.0. Incubate plate for thirty minutes while shaking.

14. Add 180 μl starvation medium to the cells and stimulate cells with 20 μl/well 10.0 mM sodium ortho vanadate and 500 ng/ml VEGF (resulting in a final

concentration of 1.0 mM sodium ortho vanadate and 50 ng/ml VEGF per well) for eight minutes at 37 °C, 5% CO_2. Negative control wells receive only starvation medium.

15. After eight minutes, media should be removed from the cells and wash one time with 200 µl/well PBS.

16. Lyse cells in 150 µl/well HNTG while shaking at room temperature for five minutes. HNTG formulation includes sodium ortho vanadate, sodium pyrophosphate and EDTA.

17. Wash ELISA plate three times as described in step 10.

18. Transfer cell lysates from the cell plate to ELISA plate and incubate while shaking for two hours. To transfer cell lysate pipette up and down while scrapping the wells.

19. Wash plate three times as described in step 10.

20. Incubate ELISA plate with 0.02 µg/well UB40 in TBSW + 05% ethanolamine. Bring final volume to 150 µl/well. Incubate while shaking for 30 minutes.

21. Wash plate three times as described in step 10.

22. Incubate ELISA plate with 1:10,000 diluted EIA grade goat anti-mouse IgG conjugated horseradish peroxidase in TBSW + 0.5% ethanolamine, pH 7.0. Bring final volume to 150 µl/well. Incubate while shaking for thirty minutes.

23. Wash plate as described in step 10.

24. Add 100 µl of ABTS/H_2O_2 solution to well. Incubate ten minutes while shaking.

25. Add 100 µl of 0.2 M HCl for 0.1 M HCl final to stop the color development reaction. Shake 1 minute at room temperature. Remove bubbles with slow stream of air and read the ELISA plate in an ELISA plate reader at 410 nm.

EXAMPLE 16: HER-2 ELISA

Assay 1: EGF Receptor-HER2 Chimeric Receptor Assay In Whole Cells.

HER2 kinase activity in whole EGFR-NIH3T3 cells was measured as described below:

Materials and Reagents

The following materials and reagents were used to conduct the assay:

1. EGF: stock concentration: 16.5 ILM; EGF 201, TOYOBO, Co., Ltd. Japan.

2. 05-101 (UBI) (a monoclonal antibody recognizing an EGFR extracellular domain).

3. Anti-phosphotyrosine antibody (anti-Ptyr) (polyclonal) (*see*, Fendley, *et al.*, *supra*).

4. Detection antibody: Goat anti-rabbit lgG horse radish peroxidase conjugate, TAGO, Inc., Burlingame, CA.

5. TBST buffer:

Tris-HCl, pH 7.2	50 mM
NaCl	150 mM
Triton X-100	0.1

6. HNTG 5X stock:

HEPES	0.1 M
NaCl	0.75 M
Glycerol	50%
Triton X-100	1.0%

7. ABTS stock:

Citric Acid	100 mM
Na_2HPO_4	250 mM
HCl, conc.	0.5 pM
ABTS*	0.5 mg/ml

* (2,2'-azinobis(3-ethylbenzthiazolinesulfonic acid)). Keep solution in dark at 4°C until use.

8. Stock reagents of:

EDTA 100 mM pH 7.0

Na_3VO_4 0.5 M

Na_4 (P_2O_7) 0.2 M

Protocol

The following protocol was used:

A. Pre-coat ELISA Plate

1. Coat ELISA plates (Corning, 96 well, Cat. #25805-96) with 05-101 antibody at 0.5 g per well in PBS, 100 µl final volume/well, and store overnight at 4 °C. Coated plates are good for up to 10 days when stored at 4 °C.

2. On day of use, remove coating buffer and replace with 100 µl blocking buffer (5% Carnation Instant Non-Fat Dry Milk in PBS). Incubate the plate, shaking, at room temperature (about 23°C to 25°C) for 30 minutes. Just prior to use, remove blocking buffer and wash plate 4 times with TBST buffer.

B. Seeding Cells

1. An NIH3T3 cell line overexpressing a chimeric receptor containing the EGFR extracellular domain and intracellular HER2 kinase domain can be used for this assay.

2. Choose dishes having 80-90% confluence for the experiment. Trypsinize cells and stop reaction by adding 10% fetal bovine serum. Suspend cells in DMEM medium (10% CS DMEM medium) and centrifuge once at 1500 rpm, at room temperature for 5 minutes.

3. Resuspend cells in seeding medium (DMEM, 0.5% bovine serum), and count the cells using trypan blue. Viability above 90% is acceptable. Seed cells in DMEM medium (0.5% bovine serum) at a density of 10,000 cells per well, 100 µl per well, in a 96 well microtiter plate. Incubate seeded cells in 5% CO_2 at 37 °C for about 4 0 hours.

C. Assay Procedures

1. Check seeded cells for contamination using an inverted microscope. Dilute drug stock (10 mg/ml in DMSO) 1:10 in DMEM medium, then transfer 5 µl to a TBST well for a final drug dilution of 1:200 and a final DMSO concentration of 1%. Control wells receive DMSO alone. Incubate in 5% CO_2 at 37 °C for two hours.

2. Prepare EGF ligand: dilute stock EGF in DMEM so that upon transfer of 10 µl dilute EGF (1:12 dilution), 100 nM final concentration is attained.

3. Prepare fresh HNTG* sufficient for 100 □l per well; and place on ice.

HNTG* (10 ml):

HNTG stock	2.0 ml

milli-Q H_2O	7.3 ml
EDTA, 100 mM, pH 7.0	0.5 ml
Na_3VO_4, 0.5 M	0.1 ml
Na_4 (P_2O_7), 0.2 M	0.1 ml

4. After 120 minutes incubation with drug, add prepared SGF ligand to cells, 10 µl per well, to a final concentration of 100 nM. Control wells receive DMEM alone. Incubate, shaking, at room temperature, for 5 minutes.

5. Remove drug, EGF, and DMEM. Wash cells twice with PBS. Transfer HNTG* to cells, 100 µl per well. Place on ice for 5 minutes. Meanwhile, remove blocking buffer from other ELISA plate and wash with TBST as described above.

6. With a pipette tip securely fitted to a micropipettor, scrape cells from plate and homogenize cell material by repeatedly aspirating and dispensing the HNTG* lysis buffer. Transfer lysate to a coated, blocked, and washed ELISA plate. Incubate shaking at room temperature for one hour.

7. Remove lysate and wash 4 times with TBST. Transfer freshly diluted anti-Ptyr antibody to ELISA plate at 100 µl per well. Incubate shaking at room temperature for 30 minutes in the presence of the anti-Ptyr antiserum (1:3000 dilution in TBST).

8. Remove the anti-Ptyr antibody and wash 4 times with TBST. Transfer the freshly diluted TAGO anti-rabbit IgG antibody to the ELISA plate at 100 µl per well. Incubate shaking at room temperature for 30 minutes (anti-rabbit IgG antibody: 1:3000 dilution in TBST).

9. Remove TAGO detection antibody and wash 4 times with TBST. Transfer freshly prepared ABTS/H_2O_2 solution to ELISA plate, 100 µl per well. Incubate shaking at room temperature for 20 minutes. (ABTS/H_2O_2 solution: 1.0 µl 30% H_2O_2 in 10 ml ABTS stock).

10. Stop reaction by adding 50 µl 5 N H_2SO_4 (optional), and determine O.D. at 4 10 nm.

11. The maximal phosphotyrosine signal is determined by subtracting the value of the negative controls from the positive controls. The percent inhibition of phosphotyrosine content for extract-containing wells is then calculated, after subtraction of the negative controls.

EXAMPLE 17: PDGF-R ELISA

All cell culture media, glutamine, and fetal bovine serum were purchased from Gibco Life Technologies (Grand Island, NY) unless otherwise specified. All cells were grown in a humid atmosphere of 90-95% air and 5-10% CO_2 at 37 °C. All cell lines were routinely subcultured twice a week and were negative for mycoplasma as determined by the Mycotect method (Gibco).

For ELISA assays, cells (U124 2, obtained from Joseph Schlessinger, NYU) were grown to 80-90% confluency in growth medium (MEM with 10% FBS, NEAA, 1 mM NaPyr and 2 mM GLN) and seeded in 96-well tissue culture plates in 0.5% serum at 25,000 to 30,000 cells per well. After overnight incubation in 0.5% serum-containing medium, cells were changed to serum-free medium and treated with test compound for 2 hr in a 5% CO_2, 37 °C incubator. Cells were then stimulated with ligand for 5-10 minute followed by lysis with HNTG (20 mM Hepes, 150 mM NaCl, 10% glycerol, 5 mM EDTA, 5 mM Na_3VO_4 , 0.2% Triton X-100, and 2 mM NaPyr). Cell lysates (0.5 mg/well in PBS) were transferred to ELISA plates previously coated with receptor-specific antibody and which had been blocked with 5% milk in TBST (50 mM Tris-HCl pH 7.2, 150 mM NaCl and 0.1% Triton X-100) at room temperature for 30 min. Lysates were incubated with shaking for 1 hour at room temperature. The plates were washed with TBST four times and then incubated with polyclonal anti-phosphotyrosine antibody at room temperature for 30 minutes. Excess anti-phosphotyrosine antibody was removed by rinsing the plate with TBST four times. Goat anti-rabbit IgG antibody was added to the ELISA plate for 30 min at room temperature followed by rinsing with TBST four more times. ABTS (100 mM citric acid, 250 mM Na_2HPO_4 and 0.5 mg/ml 2,2'-azino-bis(3-ethylbenzthiazoline-6-sulfonic acid)) plus H_2O_2 (1.2 ml 30% H_2O_2 to 10 ml ABTS) was added to the ELISA plates to start color development. Absorbance at 4 10 nm with a reference wavelength of 630 nm was recorded about 15 to 30 min after ABTS addition.

EXAMPLE 18: IGF-I Receptor ELISA

The following protocol may be used to measure phosphotyrosine level on IGF-I receptor, which indicates IGF-I receptor tyrosine kinase activity.

Materials and Reagents

The following materials and reagents were used:

1. The cell line used in this assay is 3T3/IGF-1R, a cell line genetically engineered to overexpresses IGF-1 receptor.

2. NIH3T3/IGF-1R is grown in an incubator with 5% CO_2 at 37 °C. The growth media is DMEM + 10% FBS (heat inactivated)+ 2 mM L-glutamine.

3. Affinity purified anti-IGF-1R antibody 17-69.

4. D-PBS:

KH_2PO_4	0.20 g/L
K_2HPO_4	2.16 g/L
KCl	0.20 g/L
NaCl	8.00 g/L(pH 7.2)

5. Blocking Buffer: TBST plus 5% Milk (Carnation Instant Non-Fat Dry Milk).

6. TBST buffer:

Tris-HCl	50 mM
NaCl	150 mM (pH 7.2/HCl 10 N)
Triton X-100	0.1%

Stock solution of TBS (10X) is prepared, and Triton X-100 is added to the buffer during dilution.

7. HNTG buffer:

HEPES	20 mM
NaCl	150 mM (pH 7.2/HCl 1 N)
Glycerol	10%
Triton X-100	0.2%

Stock solution (5X) is prepared and kept at 4 °C.

8. EDTA/HCl: 0.5 M pH 7.0 (NaOH) as 100X stock.

9. Na_3VO_4 : 0.5 M as 100X stock and aliquots are kept in -80 °C.

10. $Na_4P_2O_7$: 0.2 M as 100X stock.

11. Insulin-like growth factor-1 from Promega (Cat# G5111).

12. Rabbit polyclonal anti-phosphotyrosine antiserum.

13. Goat anti-rabbit IgG, POD conjugate (detection antibody), Tago (Cat. No. 4 520, Lot No. 1802): Tago, Inc., Burlingame, CA.

14. ABTS (2,2'-azinobis(3-ethylbenzthiazolinesulfonic acid)) solution:

Citric acid	100 mM
Na_2HPO_4	250 mM (pH 4.0/1 N HCl)
ABTS	0.5 mg/ml

ABTS solution should be kept in dark and 4 °C. The solution should be discarded when it turns green.

15. Hydrogen Peroxide: 30% solution is kept in the dark and at 4°C.

Protocol

All the following steps are conducted at room temperature unless it is specifically indicated. All ELISA plate washings are performed by rinsing the plate with tap water three times, followed by one TBST rinse. Pat plate dry with paper towels.

A. Cell Seeding:

1. The cells, grown in tissue culture dish (Corning 25020-100) to 80-90% confluence, are harvested with Trypsin-EDTA (0.25%, 0.5 ml/D-100, GIBCO).

2. Resuspend the cells in fresh DMEM + 10% FBS + 2 mM L-Glutamine, and transfer to 96-well tissue culture plate (Corning, 25806-96) at 20,000 cells/well (100 µl/well). Incubate for 1 day then replace medium to serum-free medium (90/µl) and incubate in 5% CO_2 and 37 °C overnight.

B. ELISA Plate Coating and Blocking:

1. Coat the ELISA plate (Corning 25805-96) with Anti-IGF-1R Antibody at 0.5 µg/well in 100 µl PBS at least 2 hours.

2. Remove the coating solution, and replace with 100 µl Blocking Buffer, and shake for 30 minutes. Remove the blocking buffer and wash the plate just before adding lysate.

C. Assay Procedures:

1. The drugs are tested in serum-free condition.

2. Dilute drug stock (in 100% DMSO) 1:10 with DMEM in 96-well poly-propylene plate, and transfer 10 µl/well of this solution to the cells to achieve final drug dilution 1:100, and final DMSO concentration of 1.0%. Incubate the cells in 5% CO_2 at 37 °C for 2 hours.

3. Prepare fresh cell lysis buffer (HNTG*)

HNTG	2 ml

EDTA	0.1 ml
Na_3VO_4	0.1 ml
$Na_4 (P_2O_7)$	0.1 ml
H_20	7.3 ml

4. After drug incubation for two hours, transfer 10 µl/well of 200nM IGF-1 Ligand in PBS to the cells (Final Conc. = 20 nM), and incubate at 5% CO_2 at 37 °C for 10 minutes.

5. Remove media and add 100 µl/well HNTG* and shake for 10 minutes. Look at cells under microscope to see if they are adequately lysed.

6. Use a 12-channel pipette to scrape the cells from the plate, and homogenize the lysate by repeated aspiration and dispensing. Transfer all the lysate to the antibody coated ELISA plate, and shake for 1 hour.

7. Remove the lysate, wash the plate, transfer anti-pTyr (1:3,000 with TBST) 100 µl/well, and shake for 30 minutes.

8. Remove anti-pTyr, wash the plate, transfer TAGO (1:3,000 with TBST) 100 µl/well, and shake for 30 minutes.

9. Remove detection antibody, wash the plate, and transfer fresh ABTS/H_2O_2 (1.2 µl H_2O_2 to 10 ml ABTS) 100 µl/well to the plate to start color development.

10. Measure OD at 4 10 nm with a reference wavelength of 630 nm in Dynatec MR5000.

EXAMPLE 19: EGF Receptor ELISA

EGF Receptor kinase activity in cells genetically engineered to express human EGF-R was measured as described below:

Materials and Reagents

The following materials and reagents were used:

1. EGF Ligand: stock concentration = 16.5 µM; EGF 201, TOYOBO, Co., Ltd. Japan.

2. 05-101 (UBI) (a monoclonal antibody recognizing an EGFR extracellular domain).

3. Anti-phosphotyosine antibody (anti-Ptyr) (polyclonal).

4. Detection antibody: Goat anti-rabbit IgG horse radish peroxidase conjugate, TAGO, Inc., Burlingame, CA.

5. TBST buffer:

Tris-HCl, pH 7	50 mM
NaCl	150 mM
Triton X-100	0.1

6. HNTG 5X stock:

HEPES	0.1 M
NaCl	0.75 M
Glycerol	50
Triton X-100	1.0%

7. ABTS stock:

Citric Acid	100 mM
Na_2HPO_4	250 mM
HCl, conc.	4.0 pH
ABTS*	0.5 mg/ml

Keep solution in dark at 4 °C until used.

8. Stock reagents of:

EDTA 100 mM pH 7.0

Na_3VO_4 0.5 M

$Na_4(P_2O_7)$ 0.2 M

Protocol

The following protocol was used:

A. Pre-coat ELISA Plate

1. Coat ELISA plates (Corning, 96 well, Cat. #25805-96) with 05-101 antibody at 0.5 µg per well in PBS, 150 µl final volume/well, and store overnight at 4 °C. Coated plates are good for up to 10 days when stored at 4 °C.

2. On day of use, remove coating buffer and replace with blocking buffer (5% Carnation Instant Non--Fat Dry Milk in PBS). Incubate the plate, shaking, at room temperature (about 23 °C to 25 °C) for 30 minutes. Just prior to use, remove blocking buffer and wash plate 4 times with TBST buffer.

B. Seeding Cells

1. NIH 3T3/C7 cell line (Honegger, *et al.*, 1987, *Cell* 51:199-209) can be use for this assay.

2. Choose dishes having 80-90% confluence for the experiment. Trypsinize cells and stop reaction by adding 10% CS DMEM medium. Suspend cells in DMEM medium (10% CS DMEM medium) and centrifuge once at 1000 rpm at room temperature for 5 minutes.

3. Resuspend cells in seeding medium (DMEM, 0.5% bovine serum), and count the cells using trypan blue. Viability above 90% is acceptable. Seed cells in DMEM medium (0.5% bovine serum) at a density of 10,000 cells per well, 100 µl per well, in a 96 well microtiter plate. Incubate seeded cells in 5% CO_2 at 37 °C for about 40 hours.

C. Assay Procedures.

1. Check seeded cells for contamination using an inverted microscope. Dilute drug stock (10 mg/ml in DMSO) 1:10 in DMEM medium, then transfer 5 µl to a test well for a final drug dilution of 1:200 and a final DMSO concentration of 1%. Control wells receive DMSO alone. Incubate in 5% CO_2 at 37 °C for one hour.

2. Prepare EGF ligand: dilute stock EGF in DMEM so that upon transfer of 10 µl dilute EGF (1:12 dilution), 25 nM final concentration is attained.

3. Prepare fresh 10 ml HNTG* sufficient for 100 µl per well wherein HNTG* comprises: HNTG stock (2.0 ml), milli-Q H_2O (7.3 ml), EDTA, 100 mM, pH 7.0 (0.5 ml), Na_3VO_4 0.5 M (0.1 ml) and $Na_4(P_2O_7)$, 0.2 M (0.1 ml).

4. Place on ice.

5. After two hours incubation with drug, add prepared EGF ligand to cells, 10 µl per well, to yield a final concentration of 25 nM. Control wells receive DMEM alone. Incubate, shaking, at room temperature, for 5 minutes.

6. Remove drug, EGF, and DMEM. Wash cells twice with PBS. Transfer HNTG* to cells, 100 µl per well. Place on ice for 5 minutes. Meanwhile, remove blocking buffer from other ELISA plate and wash with TBST as described above.

7. With a pipette tip securely fitted to a micropipettor, scrape cells from plate and homogenize cell material by repeatedly aspirating and dispensing the HNTG* lysis buffer. Transfer lysate to a coated, blocked, and washed ELISA plate. Incubate shaking at room temperature for one hour.

8. Remove lysate and wash 4 times with TBST. Transfer freshly diluted anti-Ptyr antibody to ELISA plate at 100 µl per well. Incubate shaking at room temperature for 30 minutes in the presence of the anti-Ptyr antiserum (1:3000 dilution in TBST).

9. Remove the anti-Ptyr antibody and wash 4 times with TBST. Transfer the freshly diluted TAGO 30 anti-rabbit IgG antibody to the ELISA plate at 100 µl per well. Incubate shaking at room temperature for 30 minutes (anti-rabbit IgG antibody: 1:3000 dilution in TBST).

10. Remove detection antibody and wash 4 times with TBST. Transfer freshly prepared ABTS/H_2O_2 solution to ELISA plate, 100 µl per well. Incubate at room temperature for 20 minutes. ABTS/H_2O_2 solution: 1.2 µl 30% H_2O_2 in 10 ml ABTS stock.

11. Stop reaction by adding 50 µl 5 N H_2SO_4 (optional), and determine O.D. at 410 nm.

12. The maximal phosphotyrosine signal is determined by subtracting the value of the negative controls from the positive controls. The percent inhibition of phosphotyrosine content for extract-containing wells is then calculated, after subtraction of the negative controls.

EXAMPLE 20: Met Autophosphorylation Assay – ELISA

This assay determines Met tyrosine kinase activity by analyzing Met protein tyrosine kinase levels on the Met receptor.

Materials and Reagents

The following materials and reagents were used:

1. HNTG (5X stock solution): Dissolve 23.83 g HEPES and 43.83 g NaCl in about 350 ml dH_2O. Adjust pH to 7.2 with HCl or NaOH, add 500 ml glycerol and 10 ml Triton X-100, mix, add dH_2O to 1 L total volume. To make 1 L of 1X working solution add 200 ml 5X stock solution to 800 ml dH_2O, check and adjust pH as necessary, store at 4 °C.

2. PBS (Dulbecco's Phosphate-Buffered Saline), Gibco Cat. # 450-1300EB (1X solution).

3. Blocking Buffer: in 500 ml dH_2O place 100 g BSA, 12.1 g Tris-pH7.5, 58.44 g NaCl and 10 ml Tween-20, dilute to 1 L total volume.

4. Kinase Buffer: To 500 ml dH_2O add 12.1 g TRIS pH7.2, 58.4 g NaCl, 40.7 g $MgCl_2$ and 1.9 g EGTA; bring to 1 L total volume with dH_2O.

5. PMSF (Phenylmethylsulfonyl fluoride), Sigma Cat. # P-7626, to 435.5 mg, add 100% ethanol to 25 ml total volume, vortex.

6. ATP (Bacterial Source), Sigma Cat. # A-7699, store powder at -20°C; to make up solution for use, dissolve 3.31 mg in 1 ml dH_2O.

7. RC-20H HRPO Conjugated Anti-Phosphotyrosine, Transduction Laboratories Cat. # E120H.

8. Pierce 1-Step (TM) Turbo TMB-ELISA (3,3',5,5'-tetramethylbenzidine, Pierce Cat. # 34022.

9. H_2SO_4, add 1 ml conc. (18 N) to 35 ml dH_2O.

10. Tris-HCl, Fischer Cat. # BP152-5; to 121.14 g of material, add 600 ml MilliQ H_2O, adjust pH to 7.5 (or 7.2) with HCl , bring volume to 1 L with MilliQ H_2O.

11. NaCl, Fischer Cat. # S271-10, make up 5 M solution.

12. Tween-20, Fischer Cat. # S337 -500.

13. Na_3VO_4, Fischer Cat. # S454-50, to 1.8 g material add 80 ml MilliQ H_2O, adjust pH to 10.0 with HCl or NaOH, boil in microwave, cool, check pH, repeat procedure until pH stable at 10.0, add MilliQ H_2O to 100 ml total volume, make 1 ml aliquots and store at -80°C.

14. $MgCl_2$, Fischer Cat. # M33-500, make up 1 M solution.

15. HEPES, Fischer Cat. # BP310-500, to 200 ml MilliQ H_2O, add 59.6 g material, adjust pH to 7.5, bring volume to 250 ml total, sterile filter.

16. Albumin, Bovine (BSA), Sigma Cat. # A-4503, to 30 grams material add sterile distilled water to make total volume of 300 ml, store at 4 °C.

17. TBST Buffer: to approx. 900 ml dH_2O in a 1 L graduated cylinder add 6.057 g TRIS and 8.766 g NaCl, when dissolved, adjust pH to 7.2 with HCl, add 1.0 ml Triton X-100 and bring to 1 L total volume with dH_2O.

18. Goat Affinity purified antibody Rabbit IgG (whole molecule), Cappel Cat. # 55641.

19. Anti h-Met (C-28) rabbit polyclonal IgG antibody, Santa Cruz Chemical Cat. # SC-161.

20. Transiently Transfected EGFR/Met chimeric cells (EMR) (Komada, *et al.*, 1993, *Oncogene* 8:2381-2390.

21. Sodium Carbonate Buffer, (Na_2CO_4, Fischer Cat. # S495): to 10.6 g material add 800 ml MilliQ H_2O, when dissolved adjust pH to 9.6 with NaOH, bring up to 1 L total volume with MilliQ H_2O, filter, store at 4°C.

Procedure

All of the following steps are conducted at room temperature unless it is specifically indicated otherwise. All ELISA plate washing is by rinsing 4X with TBST.

A. EMR Lysis

This procedure can be performed the night before or immediately prior to the start of receptor capture.

1. Quick thaw lysates in a 37 °C waterbath with a swirling motion until the last crystals disappear.

2. Lyse cell pellet with 1X HNTG containing 1 mM PMSF. Use 3 ml of HNTG per 15 cm dish of cells. Add ½ the calculated HNTG volume, vortex the tube for 1 min., add the remaining amount of HNTG, vortex for another min.

3. Balance tubes, centrifuge at 10,000x g for 10 min at 4°C.

4. Pool supernatants, remove an aliquot for protein determination.

5. Quick freeze pooled sample in dry ice/ethanol bath. This step is performed regardless of whether lysate will be stored overnight or used immediately following protein determination.

6. Perform protein determination using standard bicinchoninic acid (BCA) method (BCA Assay Reagent Kit from Pierce Chemical Cat. # 23225).

B. ELISA Procedure

1. Coat Corning 96 well ELISA plates with 5 µg per well Goat anti-Rabbit antibody in Carbonate Buffer for a total well volume of 50 µl. Store overnight at 4°C.

2. Remove unbound Goat anti-rabbit antibody by inverting plate to remove liquid.

3. Add 150 µl of Blocking Buffer to each well. Incubate for 30 min. at room temperature with shaking.

4. Wash 4X with TBST. Pat plate on a paper towel to remove excess liquid and bubbles.

5. Add 1 μg per well of Rabbit anti-Met antibody diluted in TBST for a total well volume of 100 μl.

6. Dilute lysate in HNTG (90 μg lysate/100 μl)

7. Add 100 μl of diluted lysate to each well. Shake at room temperature for 60 min.

8. Wash 4X with TBST. Pat on paper towel to remove excess liquid and bubbles.

9. Add 50 μl of 1X lysate buffer per well.

10. Dilute compounds/extracts 1:10 in 1X Kinase Buffer in a polypropylene 96 well plate.

11. Transfer 5.5 μl of diluted drug to ELISA plate wells. Incubate at room temperature with shaking for 20 min.

12. Add 5.5 μl of 60 μM ATP solution per well. Negative controls do not receive any ATP. Incubate at room temperature for 90 min., with shaking.

13. Wash 4X with TBST. Pat plate on paper towel to remove excess liquid and bubbles.

14. Add 100 μl per well of RC20 (1:3000 dilution in Blocking Buffer). Incubate 30 min. at room temperature with shaking.

15. Wash 4X with TBST. Pat plate on paper towel to remove excess liquid and bubbles.

16. Add 100 μl per well of Turbo-TMB. Incubate with shaking for 30-60 min.

17. Add 100 μl per well of 1 M H_2SO_4 to stop reaction.

18. Read assay on Dynatech MR7000 ELISA reader. Test Filter = 450 nm, reference filter = 410 nm.

EXAMPLE 21: Biochemical *src* Assay – ELISA

This assay is used to determine *src* protein kinase activity measuring phosphorylation of a biotinylated peptide as the readout.

Materials and Reagents

The following materials and reagents were used:

1. Yeast transformed with *src*.

2. Cell lysates: Yeast cells expressing *src* are pelleted, washed once with water, re-pelleted and stored at -80°C until use.

3. N-terminus biotinylated EEEYEEYEEEYEEEYEEEY is prepared by standard procedures well known to those skilled in the art.

4. DMSO: Sigma, St. Louis, MO.

5. 96 Well ELISA Plate: Corning 96 Well Easy Wash, Modified flat Bottom Plate, Corning Cat. #25805-96.

6. NUNC 96-well V-bottom polypropylene plates for dilution of compounds: Applied Scientific Cat. # A-72092.

7. Vecastain ELITE ABC reagent: Vector, Burlingame, CA.

8. Anti-*src* (327) mab: Schizosaccharomyces Pombe was used to express recombinant *src* (Superti-Furga, *et al.*, *EMBO J.* 12:2625-2634; Superti-Furga, *et al.*, *Nature Biochem.* 14:600-605). S. Pombe strain SP200 (h-s leul.32 ura4 ade210) was grown as described and transformations were pRSP expression plasmids were done by the lithium acetate method (Superti-Furga, *supra*). Cells were grown in the presence of 1 μM thiamin to repress expression from the nmtl promoter or in the absence of thiamin to induce expression.

9. Monoclonal anti-phosphotyrosine, UBI 05-321 (UB40 may be used instead).

10. Turbo TMB-ELISA peroxidase substrate: Pierce Chemical.

Buffer Solutions:

1. PBS (Dulbecco's Phosphate-Buffered Saline): GIBCO PBS, GIBCO Cat. # 450-1300EB.

2. Blocking Buffer: 5% Non-fat milk (Carnation) in PBS.

3. Carbonate Buffer: Na_2CO_4 from Fischer, Cat. # S495, make up 100 mM stock solution.

4. Kinase Buffer: 1.0 ml (from 1 M stock solution) $MgCl_2$; 0.2 ml (from a 1 M stock solution) $MnCl_2$; 0.2 ml (from a 1 M stock solution) DTT; 5.0 ml (from

a 1 M stock solution) HEPES; 0.1 ml TX-100; bring to 10 ml total volume with MilliQ H_2O.

5. Lysis Buffer: 5.0 HEPES (from 1 M stock solution.); 2.74 ml NaCl (from 5 M stock solution); 10 ml glycerol; 1.0 ml TX-100; 0.4 ml EDTA (from a 100 mM stock solution); 1.0 ml PMSF (from a 100 mM stock solution); 0.1 ml Na_3VO_4 (from a 0.1 M stock solution); bring to 100 ml total volume with MilliQ H_2O.

6. ATP: Sigma Cat. # A-7699, make up 10 mM stock solution (5.51 mg/ml).

7. TRIS-HCl: Fischer Cat. # BP 152-5, to 600 ml MilliQ H_2O add 121.14 g material, adjust pH to 7.5 with HCl, bring to 1 L total volume with MilliQ H_2O.

8. NaCl: Fischer Cat. # S271-10, Make up 5 M stock solution with MilliQ H_2O.

9. Na_3VO_4: Fischer Cat. # S454-50; to 80 ml MilliQ H_2O, add 1.8 g material; adjust pH to 10.0 with HCl or NaOH; boil in a microwave; cool; check pH, repeat pH adjustment until pH remains stable after heating/cooling cycle; bring to 100 ml total volume with MilliQ H_2O; make 1 ml aliquots and store at -80°C.

10. $MgCl_2$: Fischer Cat. # M33-500, make up 1 M stock solution with MilliQ H_2O.

11. HEPES: Fischer Cat. # BP 310-500; to 200 ml MilliQ H_2O, add 59.6 g material, adjust pH to 7.5, bring to 250 ml total volume with MilliQ H_2O, sterile filter (1 M stock solution).

12. TBST Buffer: TBST Buffer: To 900 ml dH_2O add 6.057 g TRIS and 8.766 g NaCl; adjust pH to 7.2 with HCl, add 1.0 ml Triton-X-100; bring to 1 L total volume with dH_2O.

13. $MnCl_2$: Fischer Cat. # M87-100, make up 1 M stock solution with MilliQ H_2O.

14. DTT; Fischer Cat. # BP172-5.

15. TBS (TRIS Buffered Saline): to 900 ml MilliQ H_2O add 6.057 g TRIS and 8.777 g NaCl; bring to 1 L total volume with MilliQ H_2O.

16. Kinase Reaction Mixture: Amount per assay plate (100 wells): 1.0 ml Kinase Buffer, 200 µg GST-ς, bring to final volume of 8.0 ml with MilliQ H_2O.

17. Biotin labeled EEEYEEYEEEYEEEYEEEY: Make peptide stock solution (1 mM, 2.98 mg/ml) in water fresh just before use.

18. Vectastain ELITE ABC reagent: To prepare 14 ml of working reagent, add 1 drop of reagent A to 15 ml TBST and invert tube several times to mix. Then add 1 drop of reagent B. Put tube on orbital shaker at room temperature and mix for 30 minutes.

Protocol

A. Preparation of *src* coated ELISA plate.

1. Coat ELISA plate with 0.5 μg/well anti-*src* mab in 100 μl of pH 9.6 sodium carbonate buffer at 4°C overnight.

2. Wash wells once with PBS.

3. Block plate with 0.15 ml 5% milk in PBS for 30 min. at room temperature.

4. Wash plate 5X with PBS.

5. Add 10 μg/well of *src* transformed yeast lysates diluted in Lysis Buffer (0.1 ml total volume per well). (Amount of lysate may vary between batches.) Shake plate for 20 minutes at room temperature.

B. Preparation of phosphotyrosine antibody-coated ELISA plate.

1. 4G10 plate: coat 0.5 μg/well 4G10 in 100 μl PBS overnight at 4°C and block with 150 μl of 5% milk in PBS for 30 minutes at room temperature.

C. Kinase assay procedure.

1. Remove unbound proteins from step 1-7, above, and wash plates 5X with PBS.

2. Add 0.08 ml Kinase Reaction Mixture per well (containing 10 μl of 10X Kinase Buffer and 10 μM (final concentration) biotin-EEEYEEYEEEYEEEYEEEY per well diluted in water.

3. Add 10 μl of compound diluted in water containing 10% DMSO and pre-incubate for 15 minutes at room temperature.

4. Start kinase reaction by adding 10 μl/well of 0.05 mM ATP in water (5 μM ATP final).

5. Shake ELISA plate for 15 min. at room temperature.

6. Stop kinase reaction by adding 10 μl of 0.5 M EDTA per well.

7. Transfer 90 μl supernatant to a blocked 4G10 coated ELISA plate from section B, above.

8. Incubate for 30 min. while shaking at room temperature.

9. Wash plate 5X with TBST.

10. Incubate with Vectastain ELITE ABC reagent (100 µl/well) for 30 min. at room temperature.

11. Wash the wells 5X with TBST.

12. Develop with Turbo TMB.

EXAMPLE 22: Biochemical *lck* Assay – ELISA

This assay is used to determine *lck* protein kinase activities measuring phosphorylation of GST-ς as the readout.

Materials and Reagents

The following materials and reagents were used:

1. Yeast transformed with *lck*. Schizosaccharomyces Pombe was used to express recombinant *lck* (Superti-Furga, *et al.*, *EMBO J.* 12:2625-2634; Superti-Furga, *et al.*, *Nature Biotech.* 14:600-605). S. Pombe strain SP200 (h-s leu1.32 ura4 ade210) was grown as described and transformations with pRSP expression plasmids were done by the lithium acetate method (Superti-Furga, *supra*). Cells were grown in the presence of 1 µM thiamin to induce expression.

2. Cell lysates: Yeast cells expressing *lck* are pelleted, washed once in water, re-pelleted and stored frozen at -80°C until use.

3. GST-ς: DNA encoding for GST-ς fusion protein for expression in bacteria obtained from Arthur Weiss of the Howard Hughes Medical Institute at the University of California, San Francisco. Transformed bacteria were grown overnight while shaking at 25°C. GST-ς was purified by glutathione affinity chromatography, Pharmacia, Alameda, CA.

4. DMSO: Sigma, St. Louis, MO.

5. 96-Well ELISA plate: Corning 96 Well Easy Wash, Modified Flat Bottom Plate, Corning Cat. #25805-96.

6. NUNC 96-well V-bottom polypropylene plates for dilution of compounds: Applied Scientific Cat. # AS-72092.

7. Purified Rabbit anti-GST antiserum: Amrad Corporation (Australia) Cat. #90001605.

8. Goat anti-Rabbit-IgG-HRP: Amersham Cat. # V010301

9. Sheep ant-mouse IgG (H+L): Jackson Labs Cat. # 5215-005-003.

10. Anti-*lck* (3A5) mab: Santa Cruz Biotechnology Cat # sc-433.

11. Monoclonal anti-phosphotyrosine UBI 05-321 (UB40 may be used instead).

Buffer solutions:

1. PBS (Dulbecco's Phosphate-Buffered Saline) 1X solution: GIBCO PBS, GIBCO Cat. # 450-1300EB.

2. Blocking Buffer: 100 g BSA, 12.1 g. TRIS-pH7.5, 58.44 g NaCl, 10 ml Tween-20, bring up to 1 L total volume with MilliQ H_2O.

3. Carbonate Buffer: Na_2CO_4 from Fischer, Cat. # S495; make up 100 mM solution with MilliQ H_2O.

4. Kinase Buffer: 1.0 ml (from 1 M stock solution) $MgCl_2$; 0.2 ml (from a 1 M stock solution) $MnCl_2$; 0.2 ml (from a 1 M stock solution) DTT; 5.0 ml (from a 1 M stock solution) HEPES; 0.1 ml TX-100; bring to 10 ml total volume with MilliQ H_2O.

5. Lysis Buffer: 5.0 HEPES (from 1 M stock solution.); 2.74 ml NaCl (from 5 M stock solution); 10 ml glycerol; 1.0 ml TX-100; 0.4 ml EDTA (from a 100 mM stock solution); 1.0 ml PMSF (from a 100 mM stock solution); 0.1 ml Na_3VO_4 (from a 0.1 M stock solution); bring to 100 ml total volume with MilliQ H_2O.

6. ATP: Sigma Cat. # A-7699, make up 10 mM stock solution (5.51 mg/ml).

7. TRIS-HCl: Fischer Cat. # BP 152-5, to 600 ml MilliQ H_2O add 121.14 g material, adjust pH to 7.5 with HCl, bring to 1 L total volume with MilliQ H_2O.

8. NaCl: Fischer Cat. # S271-10, Make up 5 M stock solution with MilliQ H_2O.

9. Na_3VO_4: Fischer Cat. # S454-50; to 80 ml MilliQ H_2O, add 1.8 g material; adjust pH to 10.0 with HCl or NaOH; boil in a microwave; cool; check pH, repeat pH adjustment until pH remains stable after heating/cooling cycle; bring to 100 ml total volume with MilliQ H_2O; make 1 ml aliquots and store at -80°C.

10. $MgCl_2$: Fischer Cat. # M33-500, make up 1 M stock solution with MilliQ H_2O.

11. HEPES: Fischer Cat. # BP 310-500; to 200 ml MilliQ H_2O, add 59.6 g material, adjust pH to 7.5, bring to 250 ml total volume with MilliQ H_2O, sterile filter (1M stock solution).

12. Albumin, Bovine (BSA), Sigma Cat. # A4503; to 150 ml MilliQ H_2O add 30 g material, bring 300 ml total volume with MilliQ H_2O, filter through 0.22 □m filter, store at 4°C.

13. TBST Buffer: To 900 ml dH_2O add 6.057 g TRIS and 8.766 g NaCl; adjust pH to 7.2 with HCl, add 1.0 ml Triton-X-100; bring to 1 L total volume with dH_2O.

14. $MnCl_2$: Fischer Cat. # M87-100, make up 1 M stock solution with MilliQ H_2O.

15. DTT; Fischer Cat. # BP172-5.

16. TBS (TRIS Buffered Saline): to 900 ml MilliQ H_2O add 6.057 g TRIS and 8.777 g NaCl; bring to 1 L total volume with MilliQ H_2O.

17. Kinase Reaction Mixture: Amount per assay plate (100 wells): 1.0 ml Kinase Buffer, 200 μg GST-ς, bring to final volume of 8.0 ml with MilliQ H_2O.

Procedures

A. Preparation of *lck* coated ELISA plate.

1. Coat 2.0 μg/well Sheep anti-mouse IgG in 100 μl of pH 9.6 sodium carbonate buffer at 4°C overnight.

2. Wash well once with PBS.

3. Block plate with 0.15 ml of blocking Buffer for 30 min. at room temp.

4. Wash plate 5X with PBS.

5. Add 0.5 μg/well of anti-*lck* (mab 3A5) in 0.1 ml PBS at room temperature for 1-2 hours.

6. Wash plate 5X with PBS.

7. Add 20 μg/well of *lck* transformed yeast lysates diluted in Lysis Buffer (0.1 ml total volume per well). (Amount of lysate may vary between batches) Shake plate at 4°C overnight to prevent loss of activity.

B. Preparation of phosphotyrosine antibody-coated ELISA plate.

1. UB40 plate: 1.0 μg/well UB40 in 100 μl of PBS overnight at 4°C and block with 150 μl of Blocking Buffer for at least 1 hour.

C. Kinase assay procedure.

1. Remove unbound proteins from step 1-7, above, and wash plates 5X with PBS.

2. Add 0.08 ml Kinase Reaction Mixture per well (containing 10 μl of 10X Kinase Buffer and 2 μg GST-ς per well diluted with water).

3. Add 10 μl of compound diluted in water containing 10% DMSO and pre-incubate for 15 minutes at room temperature.

4. Start kinase reaction by adding 10 μl/well of 0.1 mM ATP in water (10 μM ATP final).

5. Shake ELISA plate for 60 min. at room temperature.

6. Stop kinase reaction by adding 10 vl of 0.5 M EDTA per well.

7. Transfer 90 μl supernatant to a blocked 4G10 coated ELISA plate from section B, above.

8. Incubate while shaking for 30 min. at room temperature.

9. Wash plate 5X with TBST.

10. Incubate with Rabbit anti-GST antibody at 1:5000 dilution in 100 μl TBST for 30 min. at room temperature.

11. Wash the wells 5X with TBST.

12. Incubate with Goat anti-Rabbit-IgG-HRP at 1:20,000 dilution in 100 μl of TBST for 30 min. at room temperature.

13. Wash the wells 5X with TBST.

14. Develop with Turbo TMB.

EXAMPLE 23: Biochemical c-kit Assay – ELISA

A. Materials And Reagents

1) HNTG: 5X stock concentration: 100 mM HEPES pH 7.2, 750 mM NaCl, 50% glycerol, 2.5% Triton X-100.

2) PBS (Dulbecco's Phosphate-Buffered Saline): Gibco Catalog # 450-1300EB

3) 1 X Blocking Buffer: 10 mM TRIS-pH7.5, 1 % BSA, 100 mM NaCl, 0.1% Triton X-100

4) 1 X Kinase Buffer: 25 mM HEPES, 100 mM NaCl, 10 mM Mg Cl_2, 6 mM Mn Cl_2.

5) PMSF: Stock Solution = 100mM (Sigma Catalog # P-7626)

6) 10 mM ATP (Bacterial source) Sigma A-7699, 5g.

7) UB40 anti-phosphotyrosine mAb (available from Terrance at Sugen.

8) HRP conjugated sheep anti-Mouse IgG. (Amersham NA 931)

9) ABTS (5Prime-3Prime 7-579844)

10) TRIS HCL: Fisher BP 152-5

11) NaCl: Fisher S271-10

12) Triton X-100: Fisher BP151-100

13) Na_3VO_4: Fisher S454-50

14) $MgCl_2$: Fisher M33-500

15) $MnCl_2$: Fisher M87-500

16) HEPES: Fisher BP310-500

17) Albumin, Bovine (BSA): Sigma A-8551

18) TBST Buffer: 50 mM Tris pH 7.2, 150 mM NaCl, 0.1% Triton X-100.

19) Goat affinity purified antibody Rabbit IgG (whole molecule): Cappel 55641.

20) Anti Kit (C-20) rabbit polyclonal IgG antibody: Santa Cruz sc-168

21) Kit/CHO cells: CHO cells stably expressing GyrB/Kit, which are grown in standard CHO medium, supplemented with 1mg/ml G418

22) Indolinone Compounds: The indolinone compounds were synthesized as set forth in the following application: PCT application number US99/06468, filed March 26, 1999 by Fong, *et al.* and entitled METHODS OF MODULATING TYROSINE PROTEIN KINASE (Lyon & Lyon docket number 231/250 PCT which is hereby incorporated by reference in its entirety including any drawings.

B. Procedure

All of the following steps are conducted at room temperature unless it is specifically indicated. All ELISA plate washing is by rinsing 4x with TBST.

Kit Cell Lysis

This procedure is performed 1hour prior to the start of receptor capture.

1) Wash a >95% confluent 15 cm dish with PBS and aspirate as much as possible.

2) Lyse the cells with 3 ml of 1x HNTG containing 1 mM PMSF/15 cm dish. Scrape the cells from the plate and transfer to a 50 ml centrifuge tube.

3) Pool supernatants, and allow to sit, on ice, for one hour with occasional vortexing. Failure to do so with result in an increased background (approximately 3-fold higher).

4) Balance tubes and centrifuge at 10,000 x g for 10 min at 4□C. Remove an aliquot for protein determination

5) Perform protein determination as per the SOP for protein determination using the bicinchoninic acid (BCA) method.

ELISA Procedure

1) Coat Corning 96-well ELISA plates with 2 μg per well Goat anti-rabbit antibody in PBS for a total well volume of 100 μl. Store overnight at 4°C.

2) Remove unbound Goat anti-rabbit antibody by inverting plate to remove liquid.

3) Add 100 μl of Blocking Buffer to each well. Shake at room temperature for 60 min.

4) Wash 4x with TBST. Pat plate on a paper towel to remove excess liquid and bubbles

5) Add 0.2 μg per well of Rabbit anti -Kit antibody diluted in TBST for a total well volume of 100 μl. Shake at room temperature for 60 min.

6) Dilute lysate in HNTG (180 μg lysate/100 μl)

7) Add 100 μl of diluted lysate to each well. Shake at room temperature for 60 min.

8) Wash 4x with TBST. Pat plate on a paper towel to remove excess liquid and bubbles.

9) Dilute compounds/extracts (or as stated otherwise) in 1x kinase buffer, with 5μM ATP in a polypropylene 96 well plate.

10) Transfer 100 μl of diluted drug to ELISA plate wells. Incubate at room temperature with shaking for 60 min.

11) Stop reaction with the addition of 10 µl of 0.5 M EDTA. Plate is now stable for a reasonable period of time.

12) Wash 4x with TBST. Pat plate on a paper towel to remove excess liquid and bubbles.

13) Add 100 µl per well of UB40 (1:2000 dilution in TBST). Incubate 60 min at room temperature, with shaking.

14) Wash 4x with TBST. Pat plate on a paper towel to remove excess liquid and bubbles.

15) Add 100 µl per well of sheep anti-mouse IgG - HRP (1:5000 dilution in TBST). Incubate 60 min at room temperature, with shaking.

16) Wash 4x with TBST. Pat plate on a paper towel to remove excess liquid and bubbles.

17) Add 100 µl per well of ABTS. Incubate with shaking for 15-30 min.

18) Read assay on Dynatech MR7000 ELISA reader

Test Filter = 410 nm

Reference Filter = 630 nm.

EXAMPLE 24: Assay Measuring Phosphorylating Function of RAF

The following assay reports the amount of RAF-catalyzed phosphorylation of its target protein MEK as well as MEK's target MAPK. The RAF gene sequence is described in Bonner *et al.*, 1985, *Molec. Cell. Biol.* 5:1400-1407, and is readily accessible in multiple gene sequence data banks. Construction of the nucleic acid vector and cell lines utilized for this portion of the invention are fully described in Morrison *et al.*, 1988, *Proc. Natl. Acad. Sci.* USA 85:8855-8859.

Materials and Reagents

1. *Sf9 (Spodoptera frugiperda)* cells; GIBCO-BRL, Gaithersburg, MD.

2. RIPA buffer: 20 mM Tris/HCl pH 7.4, 137 mM NaCl, 10% glycerol, 1 mM PMSF, 5 mg/L Aprotenin, 0.5 % Triton X-100.

3. Thioredoxin-MEK fusion protein (T-MEK): T-MEK expression and purification by affinity chromatography were performed according to the

manufacturer's procedures. Catalog# K 350-01 and R 350-40, Invitrogen Corp., San Diego, CA.

4. His-MAPK (ERK 2); His-tagged MAPK was expressed in XL1 Blue cells transformed with pUC18 vector encoding His-MAPK. His-MAPK was purified by Ni-affinity chromatography. Cat# 27-4949-01, Pharmacia, Alameda, CA, as described herein.

5. Sheep anti mouse IgG: Jackson laboratories, West Grove, PA. Catalog, # 515-006-008, Lot# 28563.

6. RAF-1 protein kinase specific antibody: URP2653 from UBI.

7. Coating buffer: PBS; phosphate buffered saline, GIBCO-BRL, Gaithersburg, MD.

8. Wash buffer: TBST - 50 mM Tris/HCl pH 7.2, 150 mM NaCl, 0.1 % Triton X-100.

9. Block buffer: TBST, 0.1 % ethanolamine pH 7.4.

10. DMSO, Sigma, St. Louis, MO.

11. Kinase buffer (KB): 20 mM HEPES/HC1 pH 7.2, 150 mM NaCl, 0.1 % Triton X-100, 1 mM PMSF, 5 mg/L Aprotenin, 75 mM sodium ortho vanadate, 0.5 MM DTT and 10 mM $MgCl_2$.

12. ATP mix: 100 mM $MgCl_2$, 300 mM ATP, 10 mCi 33P ATP (Dupont-NEN)/ml.

13. Stop solution: 1 % phosphoric acid; Fisher, Pittsburgh, PA.

14. Wallac Cellulose Phosphate Filter mats; Wallac, Turku, Finnland.

15. Filter wash solution: 1 % phosphoric acid, Fisher, Pittsburgh, PA.

16. Tomtec plate harvester, Wallac, Turku, Finnland.

17. Wallac beta plate reader # 1205, Wallac, Turku, Finnland.

18. NUNC 96-well V bottom polypropylene plates for compounds Applied Scientific Catalog # AS-72092.

Protocol

All of the following steps were conducted at room temperature unless specifically indicated.

1. ELISA plate coating: ELISA wells are coated with 100 ml of Sheep anti mouse affinity purified antiserum (1 mg/100 ml coating buffer) over night at 4°C. ELISA plates can be used for two weeks when stored at 4°C.

2. Invert the plate and remove liquid. Add 100 ml of blocking solution and incubate for 30 min.

3. Remove blocking solution and wash four times with wash buffer. Pat the plate on a paper towel to remove excess liquid.

4. Add 1 mg of antibody specific for RAF-1 to each well and incubate for 1 hour. Wash as described in step 3.

5. Thaw lysates from RAS/RAF infected Sf9 cells and dilute with TBST to 10 mg/100 ml. Add 10 mg of diluted lysate to the wells and incubate for 1 hour. Shake the plate during incubation. Negative controls receive no lysate. Lysates from RAS/RAF infected Sf9 insect cells are prepared after cells are infected with recombinant baculoviruses at a MOI of 5 for each virus, and harvested 48 hours later. The cells are washed once with PBS and lysed in RIPA buffer. Insoluble material is removed by centrifugation (5 min at 10,000 x g). Aliquots of lysates are frozen in dry ice/ethanol and stored at -80°C until use.

6. Remove non-bound material and wash as outlined above (step 3).

7. Add 2 mg of T-MEK and 2 mg of His-MAEPK per well and adjust the volume to 40 ml with kinase buffer. Methods for purifying T-MEK and MAPK from cell extracts are provided herein by example.

8. Pre-dilute compounds (stock solution 10 mg/ml DMSO) or extracts 20 fold in TBST plus 1% DMSO. Add 5 ml of the pre-diluted compounds/extracts to the wells described in step 6. Incubate for 20 min. Controls receive no drug.

9. Start the kinase reaction by addition of 5 ml ATPmix; Shake the plates on an ELISA plate shaker during incubation.

10. Stop the kinase reaction after 60 min by addition of 30 ml stop solution to each well.

11. Place the phosphocellulose mat and the ELISA plate in the Tomtec plate harvester. Harvest and wash the filter with the filter wash solution according to the manufacturers recommendation. Dry the filter mats. Seal the filter mats and place them in the holder. Insert the holder into radioactive detection apparatus and quantify the radioactive phosphorous on the filter mats.

Alternatively, 40 ml aliquots from individual wells of the assay plate can be transferred to the corresponding positions on the phosphocellulose filter mat. After air drying the filters, put the filters in a tray. Gently rock the tray, changing the wash solution at 15 min intervals for 1 hour. Air-dry the filter mats. Seal the filter mats

and place them in a holder suitable for measuring the radioactive phosphorous in the samples. Insert the holder into a detection device and quantify the radioactive phosphorous on the filter mats.

EXAMPLE 25: CDK2/Cyclin A - Inhibition Assay

This assay analyzes the protein kinase activity of CDK2 in exogenous substrate.

Materials and Reagents

1. Buffer A (80 mM Tris (pH 7.2), 40 mM $MgCl_2$): 4.84 g Tris (F.W. =121.1 g/mol), 4.07 g $MgCl_2$ (F.W.=203.31 g/mol) dissolved in 500 ml H_2O. Adjust pH to 7.2 with HCl.

2. Histone H1 solution (0.45 mg/ml Histone H1 and 20 mM HEPES pH 7.2: 5 mg Histone H1 (Boehinger Mannheim) in 11.111 ml 20 mM HEPES pH 7.2 (477 mg HEPES (F.W.= 238.3 g/mol) dissolved in 100 ml ddH_2O), stored in 1 ml aliquots at -80°C.

3. ATP solution (60 µM ATP, 300 µg/ml BSA, 3 mM DTT): 120 µl 10 mM ATP, 600 µl 10 mg/ml BSA to 20 ml, stored in 1 ml aliquots at -80°C.

4. CDK2 solution: cdk2/cyclin A in 10 mM HEPES pH 7.2, 25 mM NaCl, 0.5 mM DTT, 10% glycerol, stored in 9 µl aliquots at -80°C.

Description of Assay:

1. Prepare solutions of inhibitors at three times the desired final assay concentration in ddH_2O/15 % DMSO by volume.

2. Dispense 20 µl of inhibitors to wells of polypropylene 96-well plates (or 20 µl 15% DMSO for positive and negative controls).

3. Thaw Histone H1 solution (1 ml/plate), ATP solution (1 ml/plate plus 1 aliquot for negative control), and CDK2 solution (9 µ/plate). Keep CDK2 on ice until use. Aliquot CDK2 solution appropriately to avoid repeated freeze-thaw cycles.

4. Dilute 9 µl CDK2 solution into 2.1 ml Buffer A (per plate). Mix. Dispense 20 µl into each well.

5. Mix 1 ml Histone H1 solution with 1 ml ATP solution (per plate) into a 10 ml screw cap tube. Add γ^{33}P ATP to a concentration of 0.15 μCi/20 μl (0.15 μCi/well in assay). Mix carefully to avoid BSA frothing. Add 20 μl to appropriate wells. Mix plates on plate shaker. For negative control, mix ATP solution with an equal amount of 20 mM HEPES pH 7.2 and add γ^{33}P ATP to a concentration of 0.15 μCi/20 μl solution. Add 20 μl to appropriate wells.

6. Let reactions proceed for 60 minutes.

7. Add 35 μl 10% TCA to each well. Mix plates on plate shaker.

8. Spot 40 μl of each sample onto P30 filter mat squares. Allow mats to dry (approx. 10-20 minutes).

9. Wash filter mats 4 X 10 minutes with 250 ml 1% phosphoric acid (10 ml phosphoric acid per liter ddH_2O).

10. Count filter mats with beta plate reader.

CELLULAR/BIOLOGIC ASSAYS

EXAMPLE 26: PDGF-Induced BrdU Incorporation Assay

Materials and Reagents:

1. PDGF: human PDGF B/B; 1276-956, Boehringer Mannheim, Germany

2. BrdU Labeling Reagent: 10 mM, in PBS (pH7.4), Cat. No. 1 647 229, Boehringer Mannheim, Germany.

3. FixDenat: fixation solution (ready to use), Cat. No. 1 647 229, Boehringer Mannheim, Germany.

4. Anti-BrdU-POD: mouse monoclonal antibody conjugated with peroxidase, Cat. No. 1 647 229, Boehringer Mannheim, Germany.

5. TMB Substrate Solution: tetramethylbenzidine (TMB), ready to use, Cat. No. 1 647 229, Boehringer Mannheim, Germany.

6. PBS Washing Solution: 1X PBS, pH 7.4, made in house (Sugen, Inc., Redwood City, California).

7. Albumin, Bovine (BSA): Fraction V powder; A-8551, Sigma Chemical Co., USA.

8. 3T3 cell line genetically engineered to express human PDGF-R.

Protocol:

1. Cells are seeded at 8000 cells/well in DMEM, 10% CS, 2 mM Gln in a 96 well plate. Cells are incubated overnight at 37 °C in 5% CO_2.

2. After 24 hours, the cells are washed with PBS, and then are serum starved in serum free medium (0% CS DMEM with 0.1% BSA) for 24 hours.

3. On day 3, ligand (PDGF, 3.8 nM, prepared in DMEM with 0.1% BSA) and test compounds are added to the cells simultaneously. The negative control wells receive serum free DMEM with 0.1% BSA only; the positive control cells receive the ligand (PDGF) but no test compound. Test compounds are prepared in serum free DMEM with ligand in a 96 well plate, and serially diluted for 7 test concentrations.

4. After 20 hours of ligand activation, diluted BrdU labeling reagent (1:100 in DMEM, 0.1% BSA) is added and the cells are incubated with BrdU (final concentration=10 μM) for 1.5 hours.

5. After incubation with labeling reagent, the medium is removed by decanting and tapping the inverted plate on a paper towel. FixDenat solution is added (50 μl/well) and the plates are incubated at room temperature for 45 minutes on a plate shaker.

6. The FixDenat solution is thoroughly removed by decanting and tapping the inverted plate on a paper towel. Milk is added (5% dehydrated milk in PBS, 200 μl/well) as a blocking solution and the plate is incubated for 30 minutes at room temperature on a plate shaker.

7. The blocking solution is removed by decanting and the wells are washed once with PBS. Anti-BrdU-POD solution (1:100 dilution in PBS, 1% BSA) is added (100 μl/well) and the plate is incubated for 90 minutes at room temperature on a plate shaker.

8. The antibody conjugate is thoroughly removed by decanting and rinsing the wells 5 times with PBS, and the plate is dried by inverting and tapping on a paper towel.

9. TMB substrate solution is added (100 µl/well) and incubated for 20 minutes at room temperature on a plate shaker until color development is sufficient for photometric detection.

10. The absorbence of the samples are measured at 410 nm (in "dual wavelength" mode with a filter reading at 490 nm, as a reference wavelength) on a Dynatech ELISA plate reader.

EXAMPLE 27: EGF-Induced BrdU Incorporation Assay

Materials and Reagents

1. EGF: mouse EGF, 201; Toyobo, Co., Ltd. Japan

2. BrdU Labeling Reagent: 10 mM, in PBS (pH7.4), Cat. No. 1 647 229, Boehringer Mannheim, Germany.

3. FixDenat: fixation solution (ready to use), Cat. No. 1 647 229, Boehringer Mannheim, Germany.

4. Anti-BrdU-POD: mouse monoclonal antibody conjugated with peroxidase, Cat. No. 1 647 229, Boehringer Mannheim, Germany.

5. TMB Substrate Solution: tetramethylbenzidine (TMB), ready to use, Cat. No. 1 647 229, Boehringer Mannheim, Germany.

6. PBS Washing Solution : 1X PBS, pH 7.4.

7. Albumin, Bovine (BSA): Fraction V powder; A-8551, Sigma Chemical Co., USA.

8. 3T3 cell line genetically engineered to express human EGF-R.

Protocol

1. Cells are seeded at 8000 cells/well in 10% CS, 2mM Gln in DMEM, in a 96 well plate. Cells are incubated overnight at 37 °C in 5% CO_2.

2. After 24 hours, the cells are washed with PBS, and then are serum starved in serum free medium (0% CS DMEM with 0.1% BSA) for 24 hours.

3. On day 3, ligand (EGF, 2 nM, prepared in DMEM with 0.1% BSA) and test compounds are added to the cells simultaneously. The negative control wells receive serum free DMEM with 0.1% BSA only; the positive control cells receive the

ligand (EGF) but no test compound. Test compounds are prepared in serum free DMEM with ligand in a 96 well plate, and serially diluted for 7 test concentrations.

4. After 20 hours of ligand activation, diluted BrdU labeling reagent (1:100 in DMEM, 0.1% BSA) is added and the cells are incubated with BrdU (final concentration=10 μM) for 1.5 hours.

5. After incubation with labeling reagent, the medium is removed by decanting and tapping the inverted plate on a paper towel. FixDenat solution is added (50 μl/well) and the plates are incubated at room temperature for 45 minutes on a plate shaker.

6. The FixDenat solution is thoroughly removed by decanting and tapping the inverted plate on a paper towel. Milk is added (5% dehydrated milk in PBS, 200 μl/well) as a blocking solution and the plate is incubated for 30 minutes at room temperature on a plate shaker.

7. The blocking solution is removed by decanting and the wells are washed once with PBS. Anti-BrdU-POD solution (1:100 dilution in PBS, 1% BSA) is added (100 μl/well) and the plate is incubated for 90 minutes at room temperature on a plate shaker.

8. The antibody conjugate is thoroughly removed by decanting and rinsing the wells 5 times with PBS, and the plate is dried by inverting and tapping on a paper towel.

9. TMB substrate solution is added (100 μl/well) and incubated for 20 minutes at room temperature on a plate shaker until color development is sufficient for photometric detection.

10. The absorbence of the samples are measured at 410 nm (in "dual wavelength" mode with a filter reading at 490 nm, as a reference wavelength) on a Dynatech ELISA plate reader.

EXAMPLE 28: EGF-Induced HER2-Driven BrdU Incorporation

Materials and Reagents:

1. EGF: mouse EGF, 201; Toyobo,Co., Ltd. Japan

2. BrdU Labeling Reagent: 10 mM, in PBS (pH 7.4), Cat. No. 1 647 229, Boehringer Mannheim, Germany.

3. FixDenat: fixation solution (ready to use), Cat. No. 1 647 229, Boehringer Mannheim, Germany.

4. Anti-BrdU-POD: mouse monoclonal antibody conjugated with peroxidase, Cat. No. 1 647 229, Boehringer Mannheim, Germany.

5. TMB Substrate Solution: tetramethylbenzidine (TMB), ready to use, Cat. No. 1 647 229, Boehringer Mannheim, Germany.

6. PBS Washing Solution : 1X PBS, pH 7.4, made in house.

7. Albumin, Bovine (BSA): Fraction V powder; A-8551, Sigma Chemical Co., USA.

8. 3T3 cell line engineered to express a chimeric receptor having the extra-cellular domain of EGF-R and the intra-cellular domain of HER2.

Protocol:

1. Cells are seeded at 8000 cells/well in DMEM, 10% CS, 2 mM Gln in a 96- well plate. Cells are incubated overnight at 37 °C in 5% CO_2.

2. After 24 hours, the cells are washed with PBS, and then are serum starved in serum free medium (0% CS DMEM with 0.1% BSA) for 24 hours.

3. On day 3, ligand (EGF=2 nM, prepared in DMEM with 0.1% BSA) and test compounds are added to the cells simultaneously. The negative control wells receive serum free DMEM with 0.1% BSA only; the positive control cells receive the ligand (EGF) but no test compound. Test compounds are prepared in serum free DMEM with ligand in a 96 well plate, and serially diluted for 7 test concentrations.

4. After 20 hours of ligand activation, diluted BrdU labeling reagent (1:100 in DMEM, 0.1% BSA) is added and the cells are incubated with BrdU (final concentration = 10 μM) for 1.5 hours.

5. After incubation with labeling reagent, the medium is removed by decanting and tapping the inverted plate on a paper towel. FixDenat solution is added (50 μl/well) and the plates are incubated at room temperature for 45 minutes on a plate shaker.

6. The FixDenat solution is thoroughly removed by decanting and tapping the inverted plate on a paper towel. Milk is added (5% dehydrated milk in PBS, 200 μl/well) as a blocking solution and the plate is incubated for 30 minutes at room temperature on a plate shaker.

7. The blocking solution is removed by decanting and the wells are washed once with PBS. Anti-BrdU-POD solution (1:100 dilution in PBS, 1% BSA) is added (100 µl/well) and the plate is incubated for 90 minutes at room temperature on a plate shaker.

8. The antibody conjugate is thoroughly removed by decanting and rinsing the wells 5 times with PBS, and the plate is dried by inverting and tapping on a paper towel.

9. TMB substrate solution is added (100 µl/well) and incubated for 20 minutes at room temperature on a plate shaker until color development is sufficient for photometric detection.

10. The absorbence of the samples are measured at 410 nm (in "dual wavelength" mode with a filter reading at 490 nm, as a reference wavelength) on a Dynatech ELISA plate reader.

EXAMPLE 29: IGF1-Induced BrdU Incorporation Assay

Materials and Reagents:

1. IGF1 Ligand: human, recombinant; G511, Promega Corp, USA.

2. BrdU Labeling Reagent: 10 mM, in PBS (pH 7.4), Cat. No. 1 647 229, Boehringer Mannheim, Germany.

3. FixDenat: fixation solution (ready to use), Cat. No. 1 647 229, Boehringer Mannheim, Germany.

4. Anti-BrdU-POD: mouse monoclonal antibody conjugated with peroxidase, Cat. No. 1 647 229, Boehringer Mannheim, Germany.

5. TMB Substrate Solution: tetramethylbenzidine (TMB), ready to use, Cat. No. 1 647 229, Boehringer Mannheim, Germany.

6. PBS Washing Solution: 1X PBS, pH 7.4.

7. Albumin, Bovine (BSA): Fraction V powder; A-8551, Sigma Chemical Co., USA.

8. 3T3 cell line genetically engineered to express human IGF-1 receptor.

Protocol:

1. Cells are seeded at 8000 cells/well in DMEM, 10% CS, 2 mM Gln in a 96- well plate. Cells are incubated overnight at 37 °C in 5% CO_2.

2. After 24 hours, the cells are washed with PBS, and then are serum starved in serum free medium (0% CS DMEM with 0.1% BSA) for 24 hours.

3. On day 3, ligand (IGF1=3.3 nM, prepared in DMEM with 0.1% BSA) and test compounds are added to the cells simultaneously. The negative control wells receive serum free DMEM with 0.1% BSA only; the positive control cells receive the ligand (IGF1) but no test compound. Test compounds are prepared in serum free DMEM with ligand in a 96 well plate, and serially diluted for 7 test concentrations.

4. After 16 hours of ligand activation, diluted BrdU labeling reagent (1:100 in DMEM, 0.1% BSA) is added and the cells are incubated with BrdU (final concentration=10 μM) for 1.5 hours.

5. After incubation with labeling reagent, the medium is removed by decanting and tapping the inverted plate on a paper towel. FixDenat solution is added (50 μl/well) and the plates are incubated at room temperature for 45 minutes on a plate shaker.

6. The FixDenat solution is thoroughly removed by decanting and tapping the inverted plate on a paper towel. Milk is added (5% dehydrated milk in PBS, 200 μl/well) as a blocking solution and the plate is incubated for 30 minutes at room temperature on a plate shaker.

7. The blocking solution is removed by decanting and the wells are washed once with PBS. Anti-BrdU-POD solution (1:100 dilution in PBS, 1% BSA) is added (100 μl/well) and the plate is incubated for 90 minutes at room temperature on a plate shaker.

8. The antibody conjugate is thoroughly removed by decanting and rinsing the wells 5 times with PBS, and the plate is dried by inverting and tapping on a paper towel.

9. TMB substrate solution is added (100 μl/well) and incubated for 20 minutes at room temperature on a plate shaker until color development is sufficient for photometric detection.

10. The absorbence of the samples are measured at 410 nm (in "dual wavelength" mode with a filter reading at 490 nm, as a reference wavelength) on a Dynatech ELISA plate reader.

EXAMPLE 30: HUV-EC-C Assay

The following protocol may also be used to measure a compound's activity against PDGF-R, FGF-R, VEGF, aFGF or Flk-1/KDR, all of which are naturally expressed by HUV-EC cells.

DAY 0

1. Wash and trypsinize HUV-EC-C cells (human umbilical vein endothelial cells, (American Type Culture Collection; catalogue no. 1730 CRL). Wash with Dulbecco's phosphate-buffered saline (D-PBS; obtained from Gibco BRL; catalogue no. 14190-029) 2 times at about 1 ml/10 cm^2 of tissue culture flask. Trypsinize with 0.05% trypsin-EDTA in non-enzymatic cell dissociation solution (Sigma Chemical Company; catalogue no. C-1544). The 0.05% trypsin was made by diluting 0.25% trypsin/1 mM EDTA (Gibco; catalogue no. 25200-049) in the cell dissociation solution. Trypsinize with about 1 ml/25-30 cm^2 of tissue culture flask for about 5 minutes at 37 °C. After cells have detached from the flask, add an equal volume of assay medium and transfer to a 50 ml sterile centrifuge tube (Fisher Scientific; catalogue no. 05-539-6).

2. Wash the cells with about 35 ml assay medium in the 50 ml sterile centrifuge tube by adding the assay medium, centrifuge for 10 minutes at approximately 200 g, aspirate the supernatant, and resuspend with 35 ml D-PBS. Repeat the wash two more times with D-PBS, resuspend the cells in about 1 ml assay medium/15 cm^2 of tissue culture flask. Assay medium consists of F12K medium (Gibco BRL; catalogue no. 21127-014) + 0.5% heat-inactivated fetal bovine serum. Count the cells with a Coulter Counter™ Coulter Electronics, Inc.) and add assay medium to the cells to obtain a concentration of 0.8-1.0x105 cells/ml.

3. Add cells to 96-well flat-bottom plates at 100 µl/well or 0.8-1.0x10^4 cells/well; incubate ~24 h at 37 °C, 5% CO2.

DAY 1

1. Make up two-fold drug titrations in separate 96-well plates, generally 50 µM on down to 0 µM. Use the same assay medium as mentioned in day 0, step 2,

above. Titrations are made by adding 90 µl/well of drug at 200 µM (4X the final well concentration) to the top well of a particular plate column. Since the stock drug concentration is usually 20 mM in DMSO, the 200 µM drug concentration contains 2% DMSO.

Therefore, diluent made up to 2% DMSO in assay medium (F12K + 0.5% fetal bovine serum) is used as diluent for the drug titrations in order to dilute the drug but keep the DMSO concentration constant. Add this diluent to the remaining wells in the column at 60 µl/well. Take 60 µl from the 120 µl of 200 µM drug dilution in the top well of the column and mix with the 60 µl in the second well of the column. Take 60 µl from this well and mix with the 60 µl in the third well of the column, and so on until two-fold titrations are completed. When the next-to-the-last well is mixed, take 60 µl of the 120 µl in this well and discard it. Leave the last well with 60 µl of DMSO/media diluent as a non-drug-containing control. Make 9 columns of titrated drug, enough for triplicate wells each for 1) VEGF (obtained from Pepro Tech Inc., catalogue no. 100-200, 2) endothelial cell growth factor (ECGF) (also known as acidic fibroblast growth factor, or aFGF) (obtained from Boehringer Mannheim Biochemica, catalogue no. 1439 600); or, 3) human PDGF B/B (1276-956, Boehringer Mannheim, Germany) and assay media control. ECGF comes as a preparation with sodium heparin.

2. Transfer 50 µl/well of the drug dilutions to the 96-well assay plates containing the 0.8-1.0x10^4 cells/100 µl/well of the HUV-EC-C cells from day 0 and incubate ~2 h at 37 °C, 5% CO_2.

3. In triplicate, add 50 µl/well of 80 µg/ml VEGF, 20 ng/ml ECGF, or media control to each drug condition. As with the drugs, the growth factor concentrations are 4X the desired final concentration. Use the assay media from day 0, step 2, to make the concentrations of growth factors. Incubate approximately 24 hours at 37 °C, 5% CO_2. Each well will have 50 µl drug dilution, 50 µl growth factor or media, and 100 µl cells, = 200 µl /well total. Thus the 4X concentrations of drugs and growth factors become 1X once everything has been added to the wells.

DAY 2

1. Add ^{3}H-thymidine (Amersham; catalogue no. TRK-686) at 1 µCi/well (10 µl/well of 100 µCi/ml solution made up in RPMI media + 10% heat-inactivated

fetal bovine serum) and incubate ~24 h at 37 °C, 5% CO_2. Note: ^{3}H-thymidine is made up in RPMI media because all of the other applications for which we use the ^{3}H-thymidine involve experiments done in RPMI. The media difference at this step is probably not significant. RPMI was obtained from Gibco BRL, catalogue no. 11875-051.

DAY 3

1. Freeze plates overnight at -20°C.

DAY 4

1. Thaw plates and harvest with a 96-well plate harvester (Tomtec Harvester 96(R)) onto filter mats (Wallac; catalogue no. 1205-401); read counts on a Wallac Betaplate(TM) liquid scintillation counter.

CONCLUSION

One skilled in the art would readily appreciate that the present invention is well adapted to carry out the objects and obtain the ends and advantages mentioned, as well as those inherent therein. The molecular complexes and the methods, procedures, treatments, molecules, specific compounds described herein are presently representative of preferred embodiments, are exemplary, and are not intended as limitations on the scope of the invention. It will be readily apparent to one skilled in the art that varying substitutions and modifications may be made to the invention disclosed herein without departing from the scope and spirit of the invention.

All patents and publications mentioned in the specification are indicative of the levels of those skilled in the art to which the invention pertains. All patents and publications are herein incorporated by reference to the same extent as if each individual publication was specifically and individually indicated to be incorporated by reference.

The invention illustratively described herein suitably may be practiced in the absence of any element or elements, limitation or limitations that are not specifically disclosed herein. Thus, for example, in each instance herein any of the terms "comprising," "consisting essentially of" and "consisting of" may be replaced with either of the other two terms. The terms and expressions which have been employed are used as terms of description and not of limitation, and there is no intention that in the use of such terms and expressions of excluding any equivalents of the features shown and described or portions thereof, but it is recognized that various modifications are possible within the scope of the invention claimed. Thus, it should be understood that although the present invention has been specifically disclosed by preferred embodiments and optional features, modification and variation of the concepts herein disclosed may be resorted to by those skilled in the art, and that such modifications and variations are considered to be within the scope of this invention as defined by the appended claims.

In addition, where features or aspects of the invention are described in terms of Markush groups, those skilled in the art will recognize that the invention is also thereby described in terms of any individual member or subgroup of members of the Markush group. For example, if X is described as selected from the group consisting

of bromine, chlorine, and iodine, claims for X being bromine and claims for X being bromine and chlorine are fully described.

In view of the degeneracy of the genetic code, other combinations of nucleic acids also encode the claimed peptides and proteins of the invention. For example, all four nucleic acid sequences GCT, GCC, GCA, and GCG encode the amino acid alanine. Therefore, if for an amino acid there exists an average of three codons, a polypeptide of 100 amino acids in length will, on average, be encoded by 3100, or 5 x 1047, nucleic acid sequences. Thus, a nucleic acid sequence can be modified to form a second nucleic acid sequence, encoding the same polypeptide as encoded by the first nucleic acid sequences, using routine procedures and without undue experimentation. Thus, all possible nucleic acids that encode the claimed peptides and proteins are also fully described herein, as if all were written out in full taking into account the codon usage, especially that preferred in humans. Furthermore, changes in the amino acid sequences of polypeptides, or in the corresponding nucleic acid sequence encoding such polypeptide, may be designed or selected to take place in an area of the sequence where the significant activity of the polypeptide remains unchanged. For example, an amino acid change may take place within a β-turn, away from the active site of the polypeptide. Also changes such as deletions (*e.g.* removal of a segment of the polypeptide, or in the corresponding nucleic acid sequence encoding such polypeptide, which does not affect the active site) and additions (*e.g.* addition of more amino acids to the polypeptide sequence without affecting the function of the active site, such as the formation of GST-fusion proteins, or additions in the corresponding nucleic acid sequence encoding such polypeptide without affecting the function of the active site) are also within the scope of the present invention. Such changes to the polypeptides can be performed by those with ordinary skill in the art using routine procedures and without undue experimentation. Thus, all possible nucleic and/or amino acid sequences that can readily be determined not to affect a significant activity of the peptide or protein of the invention are also fully described herein.

The invention has been described broadly and generically herein. Each of the narrower species and subgeneric groupings falling within the generic disclosure also form part of the invention. This includes the generic description of the invention with a proviso or negative limitation removing any subject matter from the genus, regardless of whether or not the excised material is specifically recited herein.

Other embodiments are within the following claims.

CLAIMS

What is claimed is:

1. An isolated, enriched or purified nucleic acid molecule encoding a kinase polypeptide, wherein said nucleic acid molecule comprises a nucleotide sequence that:

(a) encodes a polypeptide having an amino acid sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24;

(b) is the complement of the nucleotide sequence of (a);

(c) hybridizes under stringent conditions to the nucleotide molecule of (a) and encodes a naturally occurring kinase polypeptide;

(d) encodes a polypeptide having an amino acid sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24, except that it lacks one or more, but not all, of an N-terminal domain, a C-terminal catalytic domain, a catalytic domain, a C-terminal domain, a coiled-coil structure region, a proline-rich region, a spacer region and a C-terminal tail; or

(e) is the complement of the nucleotide sequence of (d).

2. The nucleic acid molecule of claim 1, further comprising a vector or promoter effective to initiate transcription in a host cell.

3. The nucleic acid molecule of claim 1, wherein said nucleic acid molecule is isolated, enriched, or purified from a mammal.

4. The nucleic acid molecule of claim 3, wherein said mammal is a human.

5. The nucleic acid probe of claim 1 used for the detection of nucleic acid encoding a kinase polypeptide in a sample, wherein said kinase polypeptide is

selected from the group consisting of a kinase polypeptide having an amino acid sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24.

6. A recombinant cell comprising the nucleic acid molecule of claim 1 encoding a kinase polypeptide having an amino acid sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24.

7. An isolated, enriched, or purified kinase polypeptide, wherein said polypeptide comprises an amino acid sequence having

(a) an amino acid sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24, respectively;

(b) an amino acid sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24, respectively, except that it lacks one or more, but not all, of the domains selected from the group consisting of an N-terminal domain, a C-terminal catalytic domain, a catalytic domain, a C-terminal domain, a coiled-coil structure region, a proline-rich region, a spacer region, and a C-terminal tail.

8. The kinase polypeptide of claim 7, wherein said polypeptide is isolated, purified, or enriched from a mammal.

9. The kinase polypeptide of claim 8, wherein said mammal is a human.

10. An antibody or antibody fragment having specific binding affinity to a kinase polypeptide or to a domain of said polypeptide, wherein said polypeptide is a

kinase polypeptide having an amino acid sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24.

11. A hybridoma which produces an antibody having specific binding affinity to a kinase polypeptide having an amino acid sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24.

12. A kit comprising an antibody which binds to a polypeptide of claim 7 or 8 and negative control antibody

13. A method for identifying a substance that modulates the activity of a kinase polypeptide comprising the steps of:

(a) contacting the kinase polypeptide having an amino acid sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24 with a test substance;

(b) measuring the activity of said polypeptide; and

(c) determining whether said substance modulates the activity of said polypeptide.

14. A method for identifying a substance that modulates the activity of a kinase polypeptide in a cell comprising the steps of:

(a) expressing a kinase polypeptide having an amino acid sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24;

(b) adding a test substance to said cell; and

(c) monitoring a change in cell phenotype or the interaction between said polypeptide and a natural binding partner.

15. A method for treating a disease or disorder by administering to a patient in need of such treatment a substance that modulates the activity of a kinase having an amino acid sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24.

16. The method of claim 15, wherein said disease or disorder is selected from the group consisting of cancers, immune-related diseases and disorders, cardiovascular disease, brain or neuronal-associated diseases, and metabolic disorders.

17. The method of claim 15, wherein said disease or disorder is selected from the group consisting of cancers of tissues; cancers of hematopoietic origin; diseases of the central nervous system; diseases of the peripheral nervous system; Alzheimer's disease; Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis; viral infections; infections caused by prions; infections caused by bacteria; infections caused by fungi; and ocular diseases.

18. The method of claim 15, wherein said disease or disorder is selected from the group consisting of migraines; pain; sexual dysfunction; mood disorders; attention disorders; cognition disorders; hypotension; hypertension; psychotic disorders; neurological disorders; dyskinesias; metabolic disorders; and organ transplant rejection.

19. The method of claim 15, wherein said substance modulates kinase activity *in vitro*.

20. The method of claim 19, wherein said substance is a kinase inhibitor.

21. A method for detection of a kinase polypeptide in a sample as a diagnostic tool for a disease or disorder, wherein said method comprises:

(a) contacting said sample with a nucleic acid probe which hybridizes under hybridization assay conditions to a nucleic acid target region of a kinase polypeptide having an amino acid sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24, said probe comprising the nucleic acid sequence encoding said polypeptide, fragments thereof, or the complements of said sequences and fragments; and

(b) detecting the presence or amount of the probe:target region hybrid as an indication of said disease.

22. The method of claim 21, wherein said disease or disorder is selected from the group consisting of cancers, immune-related diseases and disorders, cardiovascular disease, brain or neuronal-associated diseases, and metabolic disorders.

23. The method of claim 21, wherein said disease or disorder is selected from the group consisting of cancers of tissues; cancers of hematopoietic origin; diseases of the central nervous system; diseases of the peripheral nervous system; Alzheimer's disease; Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis; viral infections; infections caused by prions; infections caused by bacteria; infections caused by fungi; and ocular diseases.

24. The method of claim 21, wherein said disease or disorder is selected from the group consisting of migraines, pain; sexual dysfunction; mood disorders; attention disorders; cognition disorders; hypotension; hypertension; psychotic disorders; neurological disorders; dyskinesias; metabolic disorders; and organ transplant rejection.

25. A method for detection of a kinase polypeptide in a sample as a diagnostic tool for a disease or disorder, wherein said method comprises:

(a) comparing a nucleic acid target region encoding said kinase polypeptide in a sample, wherein said kinase polypeptide has an amino acid sequence selected from the group consisting of those set forth in SEQ ID NO: 13, SEQ ID NO: 14, SEQ ID NO: 15, SEQ ID NO: 16, SEQ ID NO: 17, SEQ ID NO: 18, SEQ ID NO: 19, SEQ ID NO: 20, SEQ ID NO: 21, SEQ ID NO: 22, SEQ ID NO: 23, and SEQ ID NO: 24, or one or more fragments thereof, with a control nucleic acid target region encoding said kinase polypeptide, or one or more fragments thereof; and

(b) detecting differences in sequence or amount between said target region and said control target region, as an indication of said disease or disorder.

26. The method of claim 25, wherein said disease or disorder is selected from the group consisting of cancers, immune-related diseases and disorders, cardiovascular disease, brain or neuronal-associated diseases, and metabolic disorders.

27. The method of claim 25, wherein said disease or disorder is selected from the group consisting of cancers of tissues; cancers of hematopoietic origin; diseases of the central nervous system; diseases of the peripheral nervous system; Alzheimer's disease; Parkinson's disease; multiple sclerosis; amyotrophic lateral sclerosis; viral infections; infections caused by prions; infections caused by bacteria; infections caused by fungi; and ocular diseases.

28. The method of claim 25, wherein said disease or disorder is selected from the group consisting of migraines, pain; sexual dysfunction; mood disorders; attention disorders; cognition disorders; hypotension; hypertension; psychotic disorders; neurological disorders; dyskinesias; metabolic disorders; and organ transplant rejection.

SGK216 (SEQ ID NO: 1)
GCAATGATAGTAAGAAATGCTAAAGATACAGCTCATACAAAAGCAGAATGGAATATTCTGGAGGAAGT
AAGGCATCCCTTCATTGTGGATTTAACTTATGCCTTTCAGACTGGTGGAAAACTCTACCTCATCCTTG
AGTATCTCAGTGGAGGAGAACTATTTATGCAGTTAGAAAGAGAGGGAATATTTATGGAAGACACTGCC
TGCTTTTACTTGGCAGAAATCTCCATGGCTTTGGGACATTTACATCAAAAGGGGATCATATACAGAGA
CCTGAAGCCGGAGAATATCATTCTTAATCACCAAGGTCATATGAAACTAACAGACTTTGGACTACGCA
AAGAATCTATTCATGATGGAACAGTCACACACACATTTTGTGGAACAATAGAATACATGGCCCCTGAA
ATCTTGATGAGAAGTGGCCACAATCGTGCTGTGGATTGGTGGAGTTTGGGAGCATTAATGTATGACAT
GCTGACTGGAGCACCCCCATTTACTGGGGAGAATAGAAAGAAAACAATTGACAACATCCTCAAATGTA
AACTCAATTTGCCTCCCTACCTCACACAAGAAGCCAGAGATCTGCTTAAAAAGCTGCTGAAAAGAAAT
GCTGCTTCTCTGGGAGCTGGTCCTGGGGACGCTGGAGAAGTTCAAGCTCATCCATTCTTTAGACACAT
TAACTGGGAAGAACTTCTGGCTCAAAAGGTGGAGCCCCCTTTTAAACCTCTGTTGCAATCTGAAGAGG
ATGTAAGTCAGTTTGATTCCAAGTTTACACGTCAGACACCTGTCGACAGCCCAGATGACGCAACTCTC
AGTGAAACTGCCAATCAGGTCTTTCTG

SGK237 (SEQ ID NO: 2)
ATGAGAGTATTATTTGATGAATCTGTTTTGCCACCTACAGTTTATTTTAAGAACTGCAGCATCTTGTT
CCTTGCTTCCTTGTGTGCTTTTGGTGTCCTGACTGGCTTGTTGGTTTGGTCCTTCATGCAGTATATGG
AGATTGTAGCCAATGAGTACCTCGGCTATGGAGAAGAGCAGCACACTGTGGACAAGCTGGTCAACATG
ACATATATTTTTCAAAAACTTGCTGCAGTCAAAGATCAAAGAGAATGGGTCACCACAAGTGGAGCCCA
CAAGACATTAGTAAATTTACTTGGTGCCCGAGATACTAATGTTCTATTGGGTTCCCTTCTGGCTCTGG
CTAGTTTAGCAGAAAGACTAACAGCGGAGTTGCTGCGCCTACTTTGTGCAGAGCCCCAGGTGAAAGAG
CAGGTGAAGCTCTATGAGGGGATACCGGTCCTCCTCAGTCTGCTCCACTCTGACCACTTGAAGCTCCT
CTGGAGCATTGTCTGGATTCTGGTACAGGTTTGTGAGGACCCTGAGACCAGCGTGGAAATTCGCATTT
GGGGAGGCATCAAACAGCTTCTTCATATTTTACAAGGAGACAGAAATTTTGTTTCTGATCACTCCTCC
ATTGGAAGCCTGTCCAGTGCAAATGCTGCAGGCCGAATCCAGCAGCTTCATTTATCAGAAGACTTGAG
CCCTAGGGAAATACAAGAAAATACTTTCTCACTTCAAGCAGCCTGCTGTGCTGCCCTCACTGAGCTGG
TGCTCAATGACACCAATGCCCACCAGGTGGTTCAGGAAAATGGTGTATATACAATAGCAAAATTAATT
TTACCAAATAAGCAAAAGAATGCAGCAAAAAGTAATCTATTACAGTGTTATGCTTTCAGAGCCTTGAG
ATTTCTCTTCAGTATGGAAAGAAACAGACCACTCTTTAAAAGACTTTTCCCCACAGACTTGTTTGAGA
TCTTCATTGACATAGGGCATTATGTACGTGATATCAGTGCTTATGAAGAATTGGTATCCAAGCTGAAT
TTATTAGTGGAGGATGAACTGAAGCAAATTGCTGAAAATATTGAAAGCATTAATCAGAACAAAGCTCC
TTTGAAATATATAGGCAACTATGCAATTTTGGATCATCTTGGAAGTGGAGCTTTTGGCTGTGTTTACA
AGGTTAGAAAGCATAGTGGTCAAAATCTTTTAGCAATGAAAGAGGTCAATTTACATAACCCAGCATTT
GGGAAGGATAAGAAAGATCGAGACAGCAGCGTAAGGAATATTGTTTCTGAATTAACAATAATTAAAGA
GCAGCTTTATCATCCCAACATTGTACGTTATTACAAAACATTTCTGGAAAATGATAGGTTGTACATAG
TTATGGAGCTGATAGAAGGAGCCCCGCTTGGAGAGCATTTCAGTTCTTTGAAGGAAAAACATCACCAT
TTTACTGAAGAAAGACTATGGAAAATATTTATACAGCTGTGCTTAGCTCTTCGATACTTACACAAGGA
GAAGAGGATTGTCCATAGAGATCAGACACCAAACAACATTATGTTGGGGGATAAGGACAAAGTAACCG
TTACTGACTTTGGCCTGGCAAAGCAAAAACAAGAAAACAGTAAACTCACCTCTGTGGTTGGAACAATC
CTGTATTCTTGTCCCGAGGTACTGAAGAGTGAGCCGTATGGGGAGAAGGCTGATGTCTGGGCAGTAGG
CTGCATCCTTTATCAGATGGCGACTTTGAGTCCCCCCTTCTACAGCACTAACATGCTGTCCTTGGCTA
CAAAAATAGTGGAGGCGGTATATGAACCAGTCCCAGAAGGTATCTACTCTGAAAAAGTAACAGACACC
ATCAGCAGGTGCCTCACTCCTGATGCGGAAGCTCGTCCAGATATTGTAGAAGTCAGTTCGATGATATC
AGATGTCATGATGAAATATTTAGACAACTTATCTACATCCCAGTTGTCCTTGGAAAAGAAGCTAGAAC
GGGAACGAAGACGCACACAAAGGTATTTTATGGAAGCCAACCGGAACACCGTCACATGTCACCATGAG
CTGGCTGTTCTATCTCACGAGACCTTTGAGAAGGCAAGTTTGAGTAGCAGCAGCAGTGGAGCAGCCAG
CCTGAAAAGTGAACTTTCAGAAAGCGCAGACCTGCCCCCTGAAGGCTTCCAGGCCTCCTATGGTAAAG
ACGAAGACAGGGCCTGTAACGAAATCCTGTCAGATGATAACTTCAACCTGGAAAATGCTGAGAAAGAT
ACATATTCAGAGGTAGATGATGAATTGGACATTTCGGATAACTCCAGCAGCTCCAGTTCAAGCCCTCT
GAAAGAATCTACATTCAACATTTTAAAGAGAAGTTTTAGTGCTTCAGGAGGAGAAAGACAATCCCAAA
CAAGGGACTTCACTGGAGGAACAGGATCAAGACCAAGACCAGGGCCACAGATGGGCACATTCTTGTGG
CAAGCATCAGCAGGAATTGCTGTGTCCCAGAGGAAAGTGCGTCAGATCAGTGATCCTATTCAGCAGAT
ATTAATTCAGCTGCACAAAATAATCTATATCACACAGCTTCCTCCA

Figure 1A

SGK248 (SEQ ID NO: 3)
ATGGACCATCCTAGTAGGGAAAAGGATGAAAGACAACGGACAACTAAACCCATGGCACAAAGGAGTGC
ACACTGCTCTCGACCATCTGGCTCCTCATCGTCCTCTGGGGTTCTTATGGTGGGACCCAACTTCAGGG
TTGGCAAGAAGATAGGATGTGGGAACTTCGGAGAGCTCAGATTAGGTAAAAATCTCTACACCAATGAA
TATGTAGCAATCAAACTGGAACCAATAAAATCACGTGCTCCACAGCTTCATTTAGAGTACAGATTTTA
TAAACAGCTTGGCAGTGCAGGTGAAGGTCTCCCACAGGTGTATTACTTTGGACCATGTGGGAAATATA
ATGCCATGGTGCTGGAGCTCCTTGGCCCTAGCTTGGAGGACTTGTTTGACCTCTGTGACCGAACATTT
ACTTTGAAGACGGTGTTAATGATAGCCATCCAGCTGCTTTCTCGAATGGAATACGTGCACTCAAAGAA
CCTCATTTACCGAGATGTCAAGCCAGAGAACTTCCTGATTGGTCGACAAGGCAATAAGAAAGAGCATG
TTATACACATTATAGACTTTGGACTGGCCAAGGAATACATTGACCCCGAAACCAAAAAACACATACCT
TATAGGGAACACAAAAGTTTAACTGGAACTGCAAGATATATGTCTATCAACACGCATCTTGGCAAAGA
GCAAAGCCGGAGAGATGATTTGGAAGCCCTAGGCCATATGTTCATGTATTTCCTTCGAGGCAGCCTCC
CCTGGCAAGGACTCAAGGCTGACACATTAAAAGAGAGATATCAAAAAATTGGTGACACCAAAAGGAAT
ACTCCCATTGAAGCTCTCTGTGAGAACTTTCCAGAGGAGATGGCAACCTACCTTCGATATGTCAGGCG
ACTGGACTTCTTTGAAAAACCTGATTATGAGTATTTACGGACCCTCTTCACAGACCTCTTTGAAAAGA
AAGGCTACACCTTTGACTATGCCTATGATTGGGTTGGGAGACCTATTCCTACTCCAGTAGGGTCAGTT
CACGTAGATTCTGGTGCATCTGCAATAACTCGAGAAAGCCACACACATAGGGATCGGCCATCACAACA
GCAGCCTCTTCGAAATCAGGTGGTTAGCTCAACCAATGGAGAGCTGAATGTTGATGATCCCACGGGAG
CCCACTCCAATGCACCAATCACAGCTCATGCCGAGGTGGAGGTAGTGGAGGAAGCTAAGTGCTGCTGT
TTCTTTAAGAGGAAAAGGAAGAAGACTGCTCAGCGCCACAAGTGA

SGK223 (SEQ ID NO: 4)
ATGGTCGGCCTGCACAACCTTGAGCCCCGCGGCGAGAGGAACATTGCCTTCCACCCGGTGAGCTTCCC
GGAGGAGAAGGCTGTGCACAAAGAAAAACCCTCATTTCCTTACCAAGACCGGCCCTCCACCCAGGAGA
GCTTCCGCCAGAAACTGGCTGCCTTTGCTGGGACCACATCTGGCTGTCACCAGGGCCCTGGGCCCCTG
CGGGAATCCCTGCCCTCGGAGGATGACAGTGATCAAAGGTGCTCGCCCTCCGGGGACAGCGAGGGTGG
AGAGTACTGCTCCATCCTGGACTGCTGCCCTGGGAGCCCTGTTGCCAAGGCTGCCTCCCAGACTGCAG
GTTCCCGGGGCAGGCATGGTGGCAGGGACTGCTCACCCACGTGCTGGGAGCAGGGGAAGTGTTCCGGG
CCCGCAGAGCAGGAGAAGCGGGGCCCGAGCTTCCCCAAGGAGTGCTGTAGCCAGGGCCCCACTGCCCA
CCCATCCTGCCTGGGCCCCAAGAAACTGTCCCTCACCTCGGAGGCTGCCATTTCTTCCGACGGCCTCT
CTTGTGGCAGCGGCAGCGGCAGCGGCAGCGGCGCCAGTAGCCCCTTCGTCCCCCACCTCGAGAGTGAT
TACTGCTCCCTCATGAAGGAACCTGCCCCAGAGAAGCAGCAGGACCCTGGCTGCCCAGGGGTGACCCC
TAGCAGATGCCTTGGGCTGACGGGGGAGCCCCAGCCCCCGGCCCACCCCCGGGAGGCTACACAGCCTG
AACCCATCTATGCTGAGAGCACCAAGAGGAAGAAGGCAGCTCCGGTGCCTTCCAAGTCACAGGCCAAG
ATAGAACATGCAGCTGCTGCCCAGGGCCAAGGCCAGGTATGCACAGGTAATGCCTGGGCCCAGAAAGC
AGCATCTGGCTGGGGCCGGGACAGCCCAGACCCAACTCCCCAGGTGTCAGCCACCATCACAGTCATGG
CGGCCCACCCGGAAGAGGACCATCGGACGATCTACCTGAGCAGCCCTGACTCTGCAGTGGGGGTGCAG
TGGCCACGAGGGCCTGTGAGCCAGAACTCCGAGGTAGGTGAAGAGGAGACTTCGGCTGGGCAGGGGCT
GAGCTCCAGGGAAAGCCATGCTCACAGTGCCAGCGAGAGCAAGCCCAAGGAGAGGCCCGCCATTCCCC
CCAAGTTGTCCAAGAGTAGCCCTGTAGGGTCCCCGGTGTCACCGTCTGCTGGAGGGCCCCCAGTGTCA
CCGCTGGCTGACCTTAGTGATGGGAGCTCTGGCGGCAGCAGCATTGGGCCCCAGCCTCCATCCCAAGG
TCCTGCTGACCCCGCTCCTTCCTGCCGGACCAACGGTGTCGCTATCAGTGACCCATCCAGGTGTCCCC
AGCCTGCCGCCTCGTCAGCCTCGGAACAGAGGCGGCCCAGGTTCCAGGCAGGCACCTGGAGTCGTCAG
TGCCGGATAGAGGAAGAAGAGGAGGTGGAGCAGGAATTGCTGAGTCACAGCTGGGGAAGAGAGACCAA
AAATGGCCCCACGGACCATTCAAACTCCACGACCTGGCACCGTCTCCACCCCACAGATGGCTCCTCTG
GGCAGAACAGCAAAGTTGGGACCGGGATGAGCAAATCCGCCTCTTTTGCCTTTGAGTTCCCCAAGGAC
AGAAGTGGGATTGAGACATTCTCACCTCCTCCTCCGCCTCCAAAGTCGCGGCACCTTCTAAAAATGAA
CAAGAGCAGCTCTGATTTGGAAAAAGTGAGCCAGGGCTCTGCAGAAAGCCTCAGCCCATCCTTCAGGG
GTGTCCACGTCAGCTTCACCACCGGCTCCACGGACAGCCTGGCCTCAGACTCTAGGACCTGCAGCGAT
GGAGGTCCCTCGTCTGAGCTGGCTCACTCGCCCACCAACAGCGGGAAGAAGCTCTTTGCTCCCGTTCC
GTTTCCTTCAGGCTCCACTGAGGACGTGTCCCCCAGTGGCCCCCAGCAGCCCCCTCCACTCCCCCAGA
AAAAGATAGTGAGCCGGGCAGCCTCTTCACCGGATGGCTTCTTCTGGACCCAAGGCTCCCCCAAGCCC
GGAACAGCAAGCCCCAAGCTGAACCTAAGCCACTCGGAAACCAACGTCCACGACGAATCTCACTTTAG
CTATTCGTTGAGCCCCGGGAACCGCCACCATCCTGTCTTCTCCTCTTCCGATCCTCTGGAGAAAGCTT
TCAAAGGCAGTGGCCACTGGCTTCCGGCAGCAGGGCTGGCGGGCAACAGAGGCGGCTGCGGGAGCCCT

Figure 1B

SGK223 (SEQ ID NO: 4) (CONTINUED FROM PREVIOUS PAGE)

GGCCTCCAGTGCAAAGGGGCCCCCTCCGCCTCATCCTCCCAGCTGAGCGTGTCCAGTCAAGCCTCCAC
CGGGAGCACCCAGCTTCAGCTGCACGGTCTCCTGAGCAACATCAGCAGCAAGGAGGGCACCTATGCCA
AGCTGGGGGGACTCTACACCCAGTCCCTGGCCCGCCTTGTAGCCAAATGTGAGGACCTCTTCATGGGC
GGCCAGAAAAAGGAGCTCCACTTCAATGAGAATAACTGGTCGCTCTTCAAGCTGACTTGTAACAAGCC
CTGCTGTGACTCGGGGGATGCCATTTATTACTGTGCCACCTGCTCTGAGGACCCCGGCAGCACCTATG
CTGTGAAAATCTGCAAAGCCCCTGAGCCCAAAACAGTCTCCTACTGCAGCCCGTCCGTGCCCGTGCAC
TTTAACATCCAGCAGGACTGCGGCCACTTCGTCGCCTCGGTGCCGTCCAGCATGCTCAGCTCCCCCGA
CGCGCCCAAGGACCCTGTGCCTGCCCTGCCCACACACCCCCCTGCCCAGGAGCAGGACTGCGTGGTGG
TCATCACCCGAGAGGTGCCACATCAGACCGCCTCCGACTTCGTGCGGGACTCGGCGGCCAGCCACCAG
GCGGAGCCCGAGGCGTACGAGCGGCGCGTGTGCTTCCTGCTTCTGCAACTCTGCAACGGGCTGGAGCA
CCTGAAGGAGCACGGGATCATCCACCGGGACCTGTGCCTGGAGAACCTGCTGCTGGTGCACTGCACCC
TCCAGGCCGGCCCCGGGCCCGCCCCCGCCCCGGCTCCCGCCCCCGCCGCCGCCGCGCCTCCCTGCTCC
TCTGCCGCCCCGCCTGCTGGTGGCACTCTCAGCCCCGCAGCCGGCCCCGCCTCCCCGGAAGGGCCCCG
GGAGAAGCAGCTGCCCCGGCTCATCATCAGCAACTTTTTGAAGGCCAAGCAGAAGCCGGGCGGCACCC
CAAACCTGCAGCAGAAGAAGAGCCAGGCCCGGCTGGCCCCCGAGATCGTGTCTGCTTCCCAGTACCGC
AAGTTCGATGAGTTCCAGACAGGCATCCTCATCTACGAGCTGCTGCACCAACCCAACCCGTTCGAGGT
GCGCGCCCAGCTGCGGGAGAGAGACTACCGGCAGGAGGACCTGCCGCCGCTGCCCGCGCTGTCCCTCT
ACTCACCCGGCCTGCAGCAGCTGGCACATCTGCTACTGGAGGCCGACCCCATCAAGCGTATCCGCATC
GGCGAGGCCAAGCGCGTGCTGCAGTGCCTGCTGTGGGGGCCTCGGCGCGAGCTGGTGCAGCAGCCGGG
CACCTCGGAGGAGGCGCTGTGCGGCACGCTGCACAACTGGATCGACATGAAGCGGGCCCTGATGATGA
TGAAGTTTGCGGAGAAGGCGGTGGATCGCAGGCGGGGCGTGGAGCTGGAGGACTGGCTTTGCTGCCAG
TACCTGGCGTCTGCGGAGCCCGGGGCCCTCTTACAGTCGCTGAAGCTCCTGCAGCTTCTGTGAGCCAA
GCCCCAGCCTGCACCGTCGCTGCCCCTTCCCTGCCTAACCCTTTCCTGTCTCGCCTTGGAAGCACCCA
TGTCTCCCTGGGAAATGGTACAGATGACTGGGATACCTGGATGTAAAATATATAAATATATATATAGA
AAATACATATACCATATATAAATATGAAAGACTAAGGATGCTGTTGCCCGTCCACACTCGTCTCCTCT
CTGCACTAAGTCCTCCCTGTTTTCTTCTGTAATTATACACATTTCCAGTTCCATGCAACGTCCTGAGG
ACAGTTCTGTGAACTGAATGCAGCCTGGACACTGGCCTCAATACCTTGTTTAGGATTTCTTCACCCTT
TTGTCAAATTGTTATTTAAAGAAAAAAAAAAAAGAAGAAAATTCCATTAAAATTTTTTTTGGTTGAGT
CATGAAAAAAAAAAAAAAAA

SGK269 (SEQ ID NO: 5)

ATGTCTGCTTGTAACACCTTTACTGAACATGTTTGGAAACCTGGTGAATGCAAGAATTGCTTTAAACC
TAAAAGTTTGCACCAGCTTCCCCCAGACCCTGAGAAGGCACCCATCACCCATGGCAATGTGAAAACTA
ATGCCAATCACAGTAACAACCACCGCATCAGGAACACGGGCAATTTCCGGCCTCCTGTGGCTAAAAAA
CCCACTATAGCTGTGAAGCCCACTATGATAGTGGCAGATGGGCAAAGTATATGTGGTGAGCTTAGCAT
CCAAGAACACTGTGAGAACAAACCTGTCATCATAGGGTGGAACCGAAACAGAGCTGCCTTGAGTCAGA
AACCACTTAACAATAATAATGAAGATGATGAAGGAATTAGCCATGTTCCTAAGCCTTATGGCAATAAT
GATAGTGCAAAGAAGATGTCAGATAACAATAATGGACTAACTGAAGTGTTAAAGGAGATAGCAGGCTT
GGATACTGCCCCTCAGATAAGAGGAAATGAAACAAACTCCAGAGAAACATTCTTGGGAAGAATAAATG
ATTGCTATAAACGATCATTGGAAAGAAAGCTTCCACCAAGTTGCATGATAGGTGGGATAAAGGAAACT
CAGGGCAAGCATGTTATTCTGAGTGGGAGCACAGAAGTGATTAGTAATGAAGGGGGCCGGTTCTGTTA
CCCAGAGTTTTCCAGTGGCGAGGAGAGTGAAGAGGATGTACTTTTCAGTAACATGGAGGAGGAGCACG
AGAGTTGGGATGAGAGTGATGAAGAGCTGTTGGCCATGGAGATTCGCATGAGAGGGCAACCTCGCTTT
GCCAACTTCAGAGCAAACACATTGTCTCCTGTTCGATTCTTTGTGGACAAAAAATGGAATACCATCCC
CCTGCGAAACAAGTCTCTGCAGAGAATCTGTGCTGTGGACTATGATGACAGCTATGATGAAATCCTGA
ATGGTTATGAGGAAAATTCTGTGGTCTCTTATGGACAAGGAAGCATTCAGAGCATGGTGTCATCTGAC
TCCACATCACCAGATTCTTCTTTAACAGAAGAATCACGTTCTGAGACAGCCAGTAGTTTATCCCAGAA
GATTTGTAATGGGGGATTATCTCCTGGTAACCCAGGAGATTCTAAGGACATGAAGGAAATTGAGCCCA
ATTATGAAAGTCCCTCTAGTAATAATCAGGATAAAGATTCATCACAGGCTTCCAAAAGCTCAATAAAA
GTTCCAGAGACCCACAAAGCAGTCCTTGCTCTCCGATTAGAAGAGAAAGATGGCAAGATTGCTGTACA
AACTGAGAAGGAAGAAAGTAAAGCCTCTACAGATGTTGCTGGGCAAGCAGTAACCATAAACCTTGTCC
CCACAGAAGAGCAAGCAAAACCTTACCGAGTTGTGAACCTGGAACAGCCATTGTGCAAGCCATATACT
GTCGTGGATGTGTCAGCAGCCATGGCCAGTGAGCACCTCGAGGGCCCTGTTAACAGCCCCAAGACAAA
AAGCTCATCCTCTACTCCAAACTCTCCAGTTACATCATCTTCATTGACACCAGGACAAATAAGTGCCC

Figure 1C

SGK269 (SEQ ID NO: 5) (CONTINUED FROM PREVIOUS PAGE)

ATTTCCAAAAATCCAGTGCAATTCGATACCAAGAAGTATGGACTTCTAGCACCAGTCCACGACAAAAG
ATACCTAAAGTAGAACTAATTACTAGTGGAACTGGACCAAATGTTCCTCCAAGGAAAAACTGTCACAA
ATCAGCACCTACATCACCCACAGCTACAAACATTTCCTCCAAAACCATCCCTGTTAAGTCACCTAATT
TGTCTGAAATTAAATTTAATAGTTATAACAATGCTGGTATGCCACCTTTTCCAATTATCATTCATGAC
GAGCCAACTTATGCTCGGAGTTCCAAAAATGCTATCAAAGTTCCCATTGTTATCAATCCAAATGCATA
TGACAATCTAGCTATCTACAAAAGTTTTCTGGGAACAAGTGGAGAACTCTCAGTGAAGGAAAAAACCA
CAAGTGTAATAAGCCATACTTATGAAGAAATAGAAACAGAAAGCAAAGTGCCTGATAACACCACTAGC
AAAACCACTGACTGTCTTCAAACTAAAGGGTTTTCAAACAGCACAGAGCATAAAAGGGGCTCAGTGGC
TCAGAAGGTTCAAGAGTTTAACAACTGTCTCAACAGAGGTCAGTCTTCACCACAGAGAAGCTATAGTT
CCAGCCACAGCTCCCCAGCAAAGATCCAGAGAGCCACTCAAGAGCCTGTGGCCAAAATAGAAGGCACT
CAGGAGTCTCAGATGGTGGGCAGCAGCAGCACCAGAGAGAAAGCAAGCACAGTGCTTTCTCAGATTGT
GGCTTCAATCCAACCCCCACAGTCTCCTCCAGAAACACCTCAATCTGGCCCTAAAGCTTGCAGTGTGG
AAGAGCTTTATGCCATTCCTCCAGATGCTGATGTTGCTAAGAGCACACCTAAGAGTACGCCAGTCCGG
CCCAAATCTCTCTTTACATCTCAGCCTAGTGGTGAGGCTGAAGCACCTCAGACCACAGACAGTCCTAC
CACCAAAGTACAGAAAGACCCATCCATAAAGCCAGTCACCCCCTCTCCCTCCAAATTAGTGACTAGCC
CCCAAAGTGAGCCACCAGCTCCCTTTCCCCCGCCACGCTCTACTTCTTCTCCTTACCATGCAGGTAAC
CTTTTGCAGAGGCATTTCACCAACTGGACCAAGCCAACCAGCCCTACCAGGTCAACAGAAGCTGAATC
AGTTTTGCACTCTGAAGGCAGCAGGCGGGCAGCTGATGCAAAACCTAAGCGCTGGATATCATTTAAAA
GCTTCTTCCGCCGTCGGAAAACAGATGAGGAGGATGACAAAGAGAAAGAGCGAGAGAAAGGGAAACTG
GTGGGCCTGGATGGCACAGTCATTCACATGCTGCCTCCTCCTCCAGTTCAGCGCCATCACTGGTTCAC
AGAGGCGAAAGGAGAGTCCAGTGAGAAACCAGCCATTGTCTTCATGTACAGGTGCGACCCTGCTCAAG
GCCAGCTCAGTGTGGATCAGAGCAAGGCTAGGACAGACCAGGCAGCAGTCATGGAGAAGGGTAGAGCA
GAGAATGCATTACTACAGGACTCAGAGAAGAAGAGGAGTCATTCTTCTCCATCACAGATTCCTAAAAA
GATTCTCAGTCACATGACCCATGAAGTAACAGAGGATTTTTCTCCTCGGGATCCAAGAACTGTTGTTG
GGAAGCAAGATGGCAAGGGCTGCACTTCAGTCACAACAGCATTGTCCCTACCTGAACTGGAAAGGGAA
GATGGAAAAGAAGACATTTCAGATCCTATGGACCCGAACCCTTGTAGTGCAACATACAGCAACTTAGG
GCAATCTAGAGCAGCCATGATACCTCCCAAGCAGCCACGACAGCCCAAGGGAGCTGTGGACGATGCCA
TCGCCTTTGGAGGGAAAACAGACCAAGAAGCACCCAATGCTTCCCAACCTACACCACCCCCACTGCCA
AAGAAGATGATCATAAGAGCCAATACAGAGCCAATCTCCAAGGACCTCCAAAAATCCATGGAAAGTAG
TCTTTGTGTCATGGCTAATCCCACCTATGATATCGACCCCAACTGGGATGCCAGCAGTGCTGGTTCTT
CCATCAGCTATGAACTCAAAGGACTGGACATTGAGTCTTATGACTCCTTGGAAAGGCCTTTGCGCAAG
GAGAGACCTGTCCCCTCAGCAGCAAACAGCATTTCCAGCTTAACCACTCTCAGTATTAAGGATAGATT
TTCCAACAGCATGGAATCCCTCTCCAGCCGGCGTGGGCCCTCTTGCAGACAGGGCCGAGGCATCCAGA
AGCCGCAGAGACAAGCACTTTATCGAGGACTTGAGAATCGGGAGGAAGTAGTGGGTAAAATCCGAAGC
CTTCATACAGATGCCTTGAAGAAACTGGCTGTTAAATGCGAAGACCTTTTCATGGCTGGGCAGAAAGA
CCAGCTCCGTTTTGGAGTGGACAGCTGGTCAGACTTCAGGCTAACCAGTGACAAACCATGTTGTGAGG
CAGGTGATGCGGTTTACTATACTGCTTCATATGCAAAAGATCCACTTAATAACTATGCAGTCAAGATC
TGTAAGAGCAAAGCTAAAGAATCTCAGCAGTATTATCACAGCTTGGCTGTCCGGCAGAGTCTGGCTGT
CCATTTTAACATTCAGCAGGACTGTGGTCATTTCCTTGCTGAAGTCCCTAACCGTCTGCTTCCCTGGG
AGGATCCAGATGACCCTGAAAAGGATGAGGATGACATGGAAGAGACTGAAGAAGACGCCAAAGGAGAA
ACGGATGGGAAAAACCCAAAGCCCTGTTCTGAAGCAGCATCATCCCAGAAAGAGAATCAGGGAGTCAT
GAGCAAGAAGCAGAGGAGCCACGTTGTGGTCATCACCAGGGAGGTTCCATGTCTTACTGTGGCTGATT
TTGTGCGAGACTCTCTGGCCCAGCATGGGAAAAGCCCTGATTTGTATGAGAGGCAGGTGTGTCTGCTG
CTCTTACAGCTATGCTCTGGTCTTGAGCACCTCAAACCCTACCATGTCACTCACTGCGATCTACGCCT
AGAGAACCTGCTACTTGTCCACTACCAGCCTGGGGGGACTGCCCAAGGCTTTGGGCCTGCAGAGCCCA
GCCCCACCTCATCTTATCCCACTAGGCTTATAGTGAGCAACTTCTCTCAGGCCAAGCAGAAGAGCCAT
CTGGTGGACCCCGAGATCCTCCGGGACCAGTCTCGCCTTGCCCCAGAGATCATAACAGCTACCCAGTA
TAAAAAGTGTGATGAGTTCCAGACAGGCATCCTCATCTATGAGATGCTGCACCTACCCAACCCCTTTG
ATGAGAACCCAGAGCTGAAGGAGAGGGAATACACACGAGCAGACCTGCCTCGCATCCCATTCCGCTCC
CCCTACTCCCGGGGTCTGCAGCAGCTGGCCAGCTGCCTCCTGAATCCCAACCCTTCTGAGCGGATCCT
CATTTCAGACGCCAAAGGCATCCTCCAGTGTCTGCTCTGGGGCCCCCGCGAAGATCTCTTCCAGACTT
TCACCGCCTGCCCTAGCCTAGTACAGAGGAACACCCTGCTCCAAAACTGGCTAGACATCAAGCGAACA
CTGCTCATGATCAAGTTTGCTGAGAAGTCCCTGGACAGGGAAGGTGGAATCAGCCTTGAGGACTGGCT

Figure 1D

SGK269 (SEQ ID NO: 5) (CONTINUED FROM PREVIOUS PAGE)
TTGTGCTCAGTATTTGGCTTTTGCCACTACAGACTCCCTCAGTTGTATTGTGAAAATTCTGCAGCACC
GTTAA

SGK258 (SEQ ID NO: 6)
ATGGAGACGCTTAACGGTGCCGGGGACACGGGCGGCAAGCCGTCCACGCGGGGCGGTGACCCTGCAGC
GCGGTCCCGCAGGACGGAAGGCATCCGCGCCGCGTACAGGCGGGGAGACCGCGGCGGCGCCCGGGACC
TGCTGGAGGAGGCCTGCGACCAGTGCGCGTCCCAGCTGGAAAAGGGCCAGCTTCTGAGCATCCCGGCA
GCCTATGGGGATCTGGAGATGGTCCGCTACCTACTCAGCAAGAGACTGGTGGAGCTGCCCACCGAGCC
CACGGATGACAACCCAGCCGTGGTGGCAGCGTATTTTGGACACACGGCAGTTGTGCAAAATACGCTGC
CCACCGAGCCCACGGATGACAACCCAGCCGTGGTGGCAGCGTATTTTGGACACACGGCAGTTGTGCAG
GAATTGCTTGAGTCCTTACCAGGTCCCTGCAGTCCCCAGCGGCTTCTGAACTGGATGCTGGCCTTGGC
TTGCCAGCGAGGGCACCTGGGGGTTGTGAAGCTCCTGGTCCTGACGCACGGGGCTGACCCGGAGAGCT
ACGCTGTCAGGAAGAATGAGTTCCCTGTCATCGTGCGCTTGCCCCTGTATGCGGCCATCAAGTCAGGG
AATGAAGACATTGCAATATTCCTGCTTCGGCATGGGGCCTATTTCTGTTCCTACATCTTGCTGGATAG
TCCTGACCCCAGCAAACATCTGCTGAGAAAGTACTTCATTGAAGCCAGTCCCTTGCCCAGCAGTTATC
CGGGAAAAACAGCTCTCCGTGTGAAATGGTCCCATCTCAGACTGCCCTGGGTAGACCTAGACTGGCTC
ATAGACATCTCCTGCCAGATCACGGAGCTCGACCTTTCTGCCAACTGCCTGGCGACCCTCCCCTCGGT
TATCCCCTGGGGCCTCATCAATCTCCGGAAGCTGAACCTCTCCGACAACCACCTGGGGGAGCTGCCTG
GCGTGCAGTCATCGGACGAAATCATCTGTTCCAGGCTACTTGAAATTGACATTTCCAGCAACAAGTTG
TCCCACCTCCCTCCTGGATTCTTGCACCTCTCAAAACTTCAAAAACTGACAGCTTCAAAAAATTGTTT
AGAAAAATTGTTCGAAGAAGAAAATGCCACTAACTGGATAGGTTTACGGAAGCTACAGGAACTTGATA
TATCTGACAATAAATTGACAGAACTCCCTGCCCTGTTCCTTCACTCTTTCAAGTCCCTCAATTCTCTG
AATGTCTCCAGAAACAACCTGAAGGTGTTTCCAGATCCCTGGGCCTGCCCTTTGAAATGTTGTAAAGC
TTCCAGAAATGCCCTGGAATGTCTGCCAGACAAAATGGCTGTCTTTTGGAAAAATCACCTGAAGGATG
TGGATTTCTCAGAAAACGCACTCAAAGAAGTTCCCCTGGGACTTTTCCAGCTTGATGCCCTCATGTTC
TTGAGGTTACAGGGGAACCAGCTGGCGGCACTTCCACCTCAAGAGAAGTGGACCTGCAGGCAGCTCAA
AACCCTGGATCTCTCCAGAAACCAACTTGGCAAAAATGAAGATGGACTGAAAACGAAGCGTATTGCCT
TTTTCACCACCAGAGGTCGCCAGCGCTCCGGGACTGAGGCAGAGACAACTATGGAGTTCAGTGCATCT
CTGGTAACCATTGTGTTCCTGTCTAACAACTGTAACCTCTGTGCATACACATGTGCAGCAAGTGTGCT
GGAATTTCCGGCCTTCCTAAGTGAGTCTTTGGAAGTCCTTTGCCTGAACGACAACCACCTCGACACAG
TCCCTCCCTCGGTTTGCCTACTGAAGAGCTTATCAGAGCTCTACTTGGGAAACAACCCTGGCCTCCGG
GAGCTCCCTCCTGAGCTGGGGCAGCTGGGCAACCTCTGGCAGCTGGACACTGAAGACCTGACCATCAG
CAATGTGCCTGCAGAAATCCAAAAAGAAGGCCCCAAAGCAATGCTGTCTTACCTGCGTGCTCAGCTGC
GGAAAGCGGAAAAGTGCAAGCTGATGAAGATGATCATCGTGGGTCCCCCGCGCCAGGGCAAGTCCACC
CTCCTGGAGATCTTACAGACGGGGAGGGCCCCCCAGGTGGTGCATGGAGAGGCCACCATCAGGACCAC
CAAGTGGGAGCTCCAGAGGCCGGCTGGCTCGAGAGCCAAGGTCAAGGATGGTCTGCGTGCAGAGTCCC
TGTGGGTTGAGTCCGTGGAGTTCAACGTCTGGGACATCGGGGGACCGGCCAGCATGGCCACTGTCAAC
CAGTGCTTCTTCACGGACAAGGCCCTGTACGTGGTGGTCTGGAACCTGGCGCTGGGGGAGGAGGCCGT
GGCCAACCTCCAGTTCTGGCTGCTCAACATCGAGGCCAAGGCCCCAAACGCCGTGGTGCTGGTGGTCG
GGACGCACCTGGATTTAATTGAAGCCAAGTTCCGTGTGGAAAGGATTGCAACGCTGCGTGCCTATGTG
CTGGCACTCTGCCGCTCCCCCTCCGGCTCCAGGGCCACAGGCTTCCCAGACATCACCTTCAAACACTT
ACATGAGATTTCCTGCAAGAGCCTGGAAGGTCAGGAAGGGCTGCGACAGCTGATTTTCCACGTCACGT
GCAGCATGAAGGACGTCGGCAGCACCATCGGCTGCCAGCGACTGGCAGGGCGGCTGATCCCCAGGAGC
TACCTGAGCCTGCAGGAGGCCGTGCTGGCAGAGCAGCAGCGCCGCAGCCGGGACGACGACGTGCAGTA
CCTGACGGACAGGCAGCTGGAGCAGCTGGTGGAGCAGACGCCCGACAACGACATCAAGGACTACGAGG
ACCTGCAGTCAGCCATCAGCTTCCTCATAGAAACCGGCACCCTGCTCCATTTCCCGGACACCAGCCAC
GGCCTGAGGAACCTCTACTTCCTCGACCCTATTTGGCTCTCCGAATGTCTGCAGAGGATCTTTAATAT
TAAGGGCTCTCGGTCAGTGGCCAAGAATGGGGTGATCAGAGCAGAAGACCTCAGGATGCTGCTGGTGG
GGACTGGCTTCACGCAGCAGACGGAAGAGCAGTACTTCCAGTTCCTGGCCAAGTTTGAGATCGCCCTG
CCCGTCGCCAATGACAGCTACCTCCTGCCCCATCTCCTTCCATCTAAACCTGGCCTGGACACCCACGG
TATGCGGCACCCCACAGCCAACACCATTCAGAGGGTATTTAAGATGAGCTTCGTTCCCGTTGGCTTCT
GGCAAAGGTTTATAGCACGGATGCTGATCAGCCTGGCGGAGATGGACCTGCAGCTTTTTGAAAACAAG
AAGAATACTAAAAGCAGGAACAGGAAAGTCACCATTTACAGTTTTACAGGAAACCAGAGAAATCGCTG
TAGCACATTCAGAGTGAAAAGAAATCAGACCATCTATTGGCAGGAAGGGCTCCTGGTCACTTTTGATG

Figure 1E

SGK258 (SEQ ID NO: 6) (CONTINUED FROM PREVIOUS PAGE)

GGGGCTACCTCAGTGTGGAATCTTCCGACGTGAACTGGAAAAAGAAGAAAAGCGGAGGAATGAAAATT
GTTTGCCAATCAGAAGTGAGGGACTTCTCAGCCATGGCTTTCATCACGGACCACGTCAATTCCTTGAT
TGATCAGTGGTTTCCCGCCCTGACAGCCACAGAGAGCGACGGGACGCCACTCATGGAGCAGTACGTGC
CCTGCCCGGTCTGCGAGACAGCCTGGGCCCAGCACACGGACCCCAGTGAGAAATCAGAGGATGTGCAG
TACTTCGACATGGAAGACTGTGTCCTGACGGCCATCGAGCGGGACTTCATCTCCTGCCCCAGACACCC
GGACCTCCCCGTGCCGCTGCAGGAGCTGGTCCCTGAACTGTTCATGACCGACTTCCCGGCCAGGCTCT
TCCTGGAGAACAGCAAGCTGGAGCACAGCGAGGACGAGGGCAGCGTCCTGGGCCAGGGCGGCAGTGGC
ACCGTCATCTACCGGGCCCGGTACCAGGGCCAGCCTGTGGCCGTCAAGCGCTTCCACATCAAAAAATT
CAAGAACTTTGCTAACGTACCGGCAGACACCATGCTGAGGCACCTGCGGGCCACCGATGCCATGAAGA
ACTTCTCCGAGTTCCGGCAGGAGGCCAGCATGCTGCACGCGCTGCAGCACCCCTGCATCGTGGCGCTC
ATCGGCATCAGCATCCACCCGCTCTGCTTCGCCCTGGAGCTCGCGCCGCTCAGCAGCCTCAACACCGT
GCTGTCCGAGAACGCCAGAGATTCTTCCTTTATACCCCTGGGACACATGCTCACCCAAAAAATAGCCT
ACCAGATCGCCTCGGGCCTGGCCTACCTGCACAAGAAAAACATCATCTTCTGTGACCTGAAGTCGGAC
AACATTCTGGTGTGGTCCCTTGACGTCAAGGAGCACATCAACATCAAGCTATCTGACTACGGGATTTC
GAGGCAGTCATTCCATGAGGGCGCCCTAGGCGTGGAGGGCACTCCTGGCTACCAGGCCCCAGAGATCA
GGCCTCGCATTGTATATGATGAGAAGGTAGATATGTTCTCCTATGGAATGGTGCTCTACGAGTTGCTG
TCAGGACAGCGCCCTGCACTGGGCCACCACCAGCTCCAGATTGCCAAGAAGCTGTCCAAGGGCATCCG
CCCGGTTCTGGGGCAGCCGGAGGAAGTGCAGTTCCGGCGACTGCAGGCGCTCATGATGGAGTGCTGGG
ACACTAAGCCAGAGAAGCGACCGCTGGCCCTGTCGGTGGTGAGCCAGATGAAGGACCCGACTTTTGCC
ACCTTCATGTATGAACTGTGCTGTGGGAAGCAGACAGCCTTCTTCTCATCCCAGGGCCAGGAGTACAC
CGTGGTGTTTTGGGATGGAAAAGAGGAGTCCAGGAACTACACGGTGGTGAACACAGAGAAGGGCCTCA
TGGAGGTGCAGAGGATGTGCTGCCCTGGGATGAAGGTGAGCTGCCAGCTCCAGGTCCAGAGATCCCTG
TGGACAGCCACCGAGAATTCCTACCTGGTCTTAGCGGGCCTCGCCGATGGGCTTGTGGCTGTGTTTCC
CGTGGTGCGGGGCACCCCAAAGGACAGCTGCTCCTACCTGTGCTCACACACAGCCAACAGGTCCAAGT
TCAGCATCGCGGATGAAGACGCACGGCAGAACCCCTACCCAGTGAAGGCCATGGAGGTGGTCAACAGC
GGCTCTGAGGTCTGGTACAGCAATGGGCCGGGCCTCCTTGTCATCGACTGTGCCTCCCTGGAGATCTG
CAGGCGGCTGGAGCCCTACATGGCCCCCTCCATGGTTACGTCAGTCGTGTGCAGCTCTGAGGGCAGAG
GGGAGGAGGTCGTCTGGTGCCTGGATGACAAGGCCAACTCCTTGGTGATGTACCACTCCACCACCTAC
CAGCTGTGTGCCCGGTACTTCTGCGGGGTCCCCAGCCCCCTCAGGGACATGTTTCCCGTGCGGCCCTT
GGACACGGAACCCCCGGCAGCCAGCCACACGGCCAACCCAAAGGTGCCTGAGGGGGACTCCATCGCGG
ACGTGAGCATCATGTACAGTGAGGAGCTGGGCACGCAGATCCTGATCCACCAGGAATCACTCACTGAC
TACTGCTCCATGTCCTCCTACTCCTCATCCCCACCCCGCCAGGCTGCCAGGTCCCCCTCAAGCCTCCC
CAGCTCCCCAGCAAGTTCTTCCAGTGTGCCTTTCTCCACCGACTGCGAGGACTCAGACATGCTACATA
CGCCCGGTGCTGCCTCCGACAGGTCTGAGCATGACCTGACCCCCATGGACGGGGAGACCTTCAGCCAG
CACCTGCAGGCCGTGAAGATCCTCGCCGTCAGAGACCTCATTTGGGTCCCCAGGCGCGGTGGAGATGT
TATCGTCATTGGCCTGGAGAAGGATTCTGGCGCCCAGCGGGGCCGAGTCATTGCCGTCTTAAAAGCCC
GAGAGCTGACTCCGCATGGGGTGCTGGTGGATGCTGCCGTGGTGGCAAAGGACACTGTTGTGTGCACC
TTTGAAAATGAAAACACAGAGTGGTGCCTGGCCGTCTGGAGGGGCTGGGGCGCCAGGGAGTTCGACAT
TTTCTACCAGTCCTACGAGGAGCTGGGCCGGCTGGAGGCTTGCACTCGCAAGAGAAGGTAA

SGK382 (SEQ ID NO: 7)

GCTAGTCAGGCTGGGTTCAAATCGGCCGAGGGCCAGCTGAGGGGAGGGTAGAGGGTGAAAAGCTCGGA
TCTGTGTTTGGGGAAGGCCAGGCTTGCGCTCCTCGCCGGGTTCCGCGAAGGTTAACCTTGGCTGACTT
GGCTCGCGAGCAAAGGGCAGCGTCTGAGCTCCCGGCGTTCCAGGAGTGGCCTCTTTTGTAGGAGCACC
TGAAATGCAGCGTCTGGTGCACTAAGCCGTAGCGGCAGCAGCAGCCACAGCGACAGCGCTGGGGCCCT
GTGTAGAAGCTCCATCCCCCTTGTCTTTGTGCTTGCCTGCGTCCCCAGACTCAGAGATTATCTTAGAA
GACCTAGGACTCCAAAAATGTTTCCCCTGAAGGACGCTGAAATGGGAGCCTTTACCTTCTTTGCCTCG
GCTCTGCCACATGATGTTTGTGGAAGCAATGGACTTCCTCTCACACCAAATTCCATCAAAATTTTAGG
GCGCTTTCAAATCCTTAAAACCATCACCCATCCCAGACTCTGCCAGTATGTGGATATTTCTAGGGGAA
AGCATGAACGACTAGTGGTCGTGGCTGAACATTGTGAACGTAGTCTGGAAGACTTGCTTCGAGAAAGG
AAACCTGTGAGCTGTTCAACGGTTTTGTGTATAGCATTTGAGGTTCTTCAGGGCTTGCAGTATATGAA
CAAACATGGTATAGTACACAGGGCATTGTCTCCTCATAATATCCTGTTGGACCGAAAGGGACATATTA
AATTGGCTAAATTTGGACTTTATCACATGACAGCTCATGGTGATGATGTTGATTTCCCAATAGGGTAT
CCCTCGTACTTGGCCCCTGAGGTAATTGCACAGGGAATTTTCAAAACCACTGATCACATGCCAAGTAA

SGK382 (SEQ ID NO: 7) (CONTINUED FROM PREVIOUS PAGE)
AAAACCATTGCCTTCTGGCCCCAAATCAGATGTATGGTCTCTTGGAATCATTTTATTTGAGCTTTGTG
TGGGAAGAAAATTATTTCAGAGCTTGGATATTTCTGAAAGACTAAAATTTTTGCTTACTTTGGATTGT
GTAGATGACACTTTAATAGTTCTGGCTGAAGAGCATGGTTGTTTGGACATTATAAAGGAGCTTCCTGA
AACTGTGATAGATCTTTTGAATAAGTGCCTTACCTTCCATCCTTCTAAGAGGCCAACCCCAGATGAAT
TAATGAAGGACAAAGTATTCAGTGAGGTATCACCTTTATATACCCCCTTTACCAAACCTGCCAGTCTG
TTTTCATCTTCTCTGAGATGTGCTGATTTAACTCTGCCTGAGGATATCAGTCAGTTGTGTAAAGATAT
AAATAATGATTACCTGGCAGAAAGATCTATTGAAGAAGTGTATTACCTTTGGTGTTTGGCTGGAGGTG
ACTTGGAGAAAGAGCTTGTCAACAAGGAAATCATTCGATCCAAACCACCTATCTGCACACTCCCCAAT
TTTCTCTTTGAGGATGGTGAAAGCTTTGGACAAGGTCGAGATAGAAGCTCGCTTTTAGATGATACCAC
TGTGACATTGTCGTTATGCCAGCTAAGAAATAGATTGAAAGATGTTGGTGGAGAAGCATTTTACCCAT
TACTTGAAGATGACCAGTCTAATTTACCTCATTCAAACAGCAATAATGAGTTGTCTGCAGCTGCCACG
CTCCCTTTAATCATCAGAGAGAAGGATACAGAGTACCAACTAAATAGAATTATTCTCTTCGACAGGCT
GCTAAAGGCTTATCCATATAAAAAAAACCAAATCTGGAAAGAAGCAAGAGTTGACATTCCTCCTCTTA
TGAGAGGTTTAACCTGGGCTGCTCTTCTGGGAGTTGAGGGAGCTATTCATGCCAAGTACGATGCAATT
GATAAAGACACTCCAATTCCTACAGATAGACAAATTGAAGTGGATATTCCTCGCTGTCATCAGTACGA
TGAACTGTTATCATCACCAGAAGGTCATGCAAAATTTAGGCGTGTATTAAAAGCCTGGGTAGTGTCTC
ATCCTGATCTTGTGTATTGGCAAGGTCTTGACTCACTTTGTGCTCCATTCCTATATCTAAACTTCAAT
AATGAAGCCTTGGCTTATGCATGTATGTCTGCTTTTATTCCCAAATACCTGTATAACTTCTTCTTAAA
AGACAACTCACATGTAATACAAGAGTATCTGACTGTCTTCTCTCAGATGATTGCATTTCATGATCCAG
AGCTGAGTAATCATCTCAATGAGATTGGTTTCATTCCAGATCTCTATGCCATCCCTTGGTTTCTTACC
ATGTTTACTCATGTATTTCCACTACACAAAATTTTCCACCTCTGGGATACCTTACTACTTGGGAATTC
CTCTTTCCCATTCTGTATTGGAGTAGCAATTCTTCAGCAGCTGCGGGACCGGCTTTTGGCTAATGGCT
TTAATGAGTGTATTCTTCTCTTCTCCGATTTACCAGAAATTGACATTGAACGCTGTGTGAGAGAATCT
ATCAACCTGTTTTGTTGGACTCCTAAAAGTGCTACTTACAGACAGCATGCTCAACCTCCAAAGCCATC
TTCTGACAGCAGTGGAGGCAGAAGTTCGGCACCTTATTTCTCTGCTGAGTGTCCAGATCCTCCAAAGA
CAGATCTGTCAAGAGAATCCATCCCATTAAATGACCTGAAGTCAGAAGTATCACCACGGATTTCAGCA
GAGGACCTGATTGACTTGTGTGAGCTCACAGTGACAGGCCACTTCAAAACACCCAGCAAGAAAACAAA
GTCCAGTAAACCAAAGCTCCTGGTGGTTGACATCCGGAATAGTGAAGACTTTATTCGTGGTCACATTT
CAGGAAGCATCAACATTCCATTCAGTGCTGCCTTCACTGCAGAAGGGGAGCTTACCCAGGGCCCTTAC
ACTGCTATGCTCCAGAACTTCAAAGGGAAGGTCATTGTCATCGTGGGGCATGTGGCAAAACACACAGC
TGAGTTTGCAGCTCACCTTGTGAAGATGAAATATCCAAGAATCTGTATTCTAGATGGTGGCATTAATA
AAATAAAGCCAACAGGCCTCCTCACCATCCCATCTCCTCAAATATGAAGAACCAAGAGTGTGACTGCC
AAAACTTAGTGTGGCATCAGCACCAACAGCACAGTTCTTCATATCCACGCCACTCTCAGACAAAACTA
GATGTCCAGATTGTTGCATTTCCGTAAAGTTTGTCACGAGACATTTTTTAAAATCTCATAACCCACAT
GTTCAGTTATCCATGCAAGAAACTTGACTCTACATGTATTGCTGAAAGAATTTTCTTAACAGTGAAAT
CTGATCATATATTTTTACCACACTGCCACATAAAGCCCAAGAAATTCAGCTGACAAGACAGATTTAGC
ATTATCAAGAAATCCCATTTGCCCTGAAAAAGCTGTCCTCCATTGTACTGAACAGACAGTCCTGTCGA
TTGTGTTATTTAGAAACATACACTGAATGTGGGCTGAAATCATCATCTTTCCATAATGAAAACTGAGA
AACTATTCACAATGCATTCCTTATAAATAAATGCTACATTTAGTAACTCATTTCACCCAAACAAGGGG
GATGTGTGTGTGTGTGTGTGTATAGGAAGTGGAGTTTATCCCATTGCAGAAAGTGGTAATACTTACTC
CCAGAAAAATGAAATTTAGAAACCATTTATATTTGATAGAATATTTGGTCAGTTCCTGTAGCAAAGAC
GAATGGCTTAAACAAATTTTCTAGTTTCTTTATCACATGAAAGTCTGTACAGTCAGTCCAGGGCTAGT
CTACTGGTTTCCTGATCATTAAGAACTCATTACCTTCTCTCATTGCTTTACAAACCTCAATATGTGGC
ATCCATCTCATGGATGAAAATGGCTCCTCAGCTTCTACCATCACATCTGCTATCTAGAAGGAAGAGAA
TGAGGGAAGGAGGGAGGGGATGAAGAGAAAAGAAGGAAC

SGK424 (SEQ ID NO: 8)
ATGGAGCCGACTGAGCCGCAGCCCACAGATGCCCTCGCAGCCCTTTCGCGCTTCCCAGTGGAGCTGGG
CTCCAAGAGGGATCTAGACACCCAGGTGGGCCGTCTGCACCACCTGGCTGTGGCTGGGGGCCCAGGAT
GGGGCCCGGCCAGGCTCCTGCGGAGAGGTGAGGGCCTGGCCGCAGGTGTCCAGGTGGACGGGGAGAAC
TCTGAGGGGCAGACAGCGCTCTTCCTCTCGGCGCTGCTGGGCCACGGCTCTGCCATGCAGCTCCTGCT
GGCCTTTGGTGCCAAGCCCAACCAGTGGCTGGCCAGCCATCCCTTCCCTAATCGACTCCCACCGCGTT
GTTCTGCTCTAACCGCTGCCTGGATGGCAGCACCCCCGTGCCTGCAGGCGCCTTCTTTAGTGATGGTG

Figure 1G

SGK424 (SEQ ID NO: 8) (CONTINUED FROM PREVIOUS PAGE)

CCCCTGCTGCAGGCAGGGGGTGACTTGCAACTGCATGACCAGCGGGGTCGCACACCACAAGACTGGGC
TGAGCAGGGTGGGGCCAAGCAGAACTGGGAGTTGCTGGAGTTGTTGCAGCTCTGCCGGGCCCACATAT
CAGCGTTCGTGCATGACGGTGAGTTGCCACCCATGGCATCTCTGGACCAGTTGCAGGCCAGATTTGGG
AACAGCCCGCCTGGCTCCCTGTCCTCCCTGAGGCTGATGCAGGCAGACAGGATGCTGAGGCTAGCACA
GATCAGGAGAGCCTCCCAAGTCCCGGCCTTGGGGTTTGGTCAGCTGAGTAGCCTGTGGCCACCGGGGC
TGGTAACGGGCATCCCCCTTGTGGACCCCAAGGAGCTGCTGGCAGCCCAGGGAGAGCCTGACTGCACC
TATGAGAGCAGCTCCCCACCCTCATGGCCAAGGAAAGCCCCCTCGTCTCAGGCCTGTACTGCCCCGCC
CCTATTCTCCAGCCTCCTGTGGATGGGTCACAAAATGACTGTGCAGCAGCCGAAGGCCCCTGGAACCC
AGACAGATGTGCTACTGGCTGACCTTCAGCACTGCAGCAAGCTGCACCACCCTGGCCTGTTGCTGCTG
ATGGCACTGAGTCCCTCTGCAGATCTGTCGGGGCTGTGCCTTCTCTTTGAGCCTGTGTGGCTGGGCTC
CCTGCAGGGGCTGGTGCAGCCAGGCCTGGCTAAAGTGGGCCACCTGGAGCACAGGCGCCTCCTACGCC
AGCGCTGGCTGCGGCCCAGGCAGCAGAAGGGCTACCCCTGGGGAGGCCCAGGCCCAGGGCTTCCCCCG
CCCCCTGAACTACACCCATGGCTGCCACTTCAGCTGATCTGCGGTGACATGCCCACCACCACCTCAGA
CCTCTACAGCTTCTGCATCCTGATCCAGGAGGTCTTCACTGGAGAGCTGCCCTGGGCTGGGAGAGGGG
GACCTGAGGTGAAGGCCAAGCTGGAGGCAGGGGAGAGTCCGGCCCTGGAGCCCCGGGTGCCAGCACCC
TACCAGGCCCTGATTCGAGCTGGGCTAGGCCTAGCACCTGCTGACCGCTGGGGCAGCCTGCAGAGTAC
CCGATACCTACTGCGGGAGGCCCCGGCCCAGAGAAACTGTACATCAAGTGGCTCACAGAGCCAAGACC
ACACCCAGGCCTGTGGCCCCTGTGATGTCCCCAGGCCCCATCCCATGCCAGGCCCGCTGGAGCACCAT
GGCTACCAGGTGGCAGCTCCAGCCTTCTCTCCTGCAGGCACCCGGTACTGA

SGK251 (SEQ ID NO: 9)

CAATTCCCCTCGCCTCCAGCCGCGAGGAGCTCCCCGGCGCCGCAGGCAGCGTCCTCCTCCGAAGCAGC
TGCACCTGCAACTGGGCAGCCTGGACCCTCGTGCCCTGTTCCCGGGACCTCGCGCAGGGGGCGCCCCG
GGACACCCCCTGCGGGCCGGGTGGAGGAGGAAGAGGAGGAGGAGGAAGAAGACGTGGACAAGGACCCC
CATCCTACCCAGAACACCTGCCTGCGCTGCCGCCACTTCTCTTTAAGGGAGAGGAAAAGAGAGCCTAG
GAGAACCATGGGGGGCTGCGAAGTCCGGGAATTTCTTTTGCAATTTGGTTTCTTCTTGCCTCTGCTGA
CAGCGTGGCCAGGCGACTGCAGTCACGTCTCCAACAACCAAGTTGTGTTGCTTGATACAACAACTGTA
CTGGGAGAGCTAGGATGGAAAACATATCCATTAAATGGGTGGGATGCCATCACTGAAATGGATGAACA
TAATAGGCCCATTCACACATACCAGGTATGTAATGTAATGGAACCAAACCAAAACAACTGGCTTCGTA
CAAACTGGATCTCCCGTGATGCAGCTCAGAAAATTTATGTGGAAATGAAATTCACACTAAGGGATTGT
AACAGCATCCCATGGGTCTTGGGGACTTGCAAAGAAACATTTAATCTGTTTTATATGGAATCAGATGA
GTCCCACGGAATTAAATTCAAGCCAAACCAGTATACAAAGATCGACACAATTGCTGCTGATGAGAGTT
TTACCCAGATGGATTTGGGTGATCGCATCCTCAAACTCAACACTGAAATTCGTGAGGTGGGGCCTATA
GAAAGGAAAGGATTTTATCTGGCTTTTCAAGACATTGGGGCGTGCATTGCCCTGGTTTCAGTCCGTGT
TTTCTACAAGAAATGCCCCTTCACTGTTCGTAACTTGGCCATGTTTCCTGATACCATTCCAAGGGTTG
ATTCCTCCTCTTTGGTTGAAGTACGGGGTTCTTGTGTGAAGAGTGCTGAAGAGCGTGACACTCCTAAA
CTGTATTGTGGAGCTGATGGAGATTGGCTGGTTCCTCTTGGAAGGTGCATCTGCAGTACAGGATATGA
AGAAATTGAGGGTTCTTGCCATGCTTGCAGACCAGGATTCTATAAAGCTTTTGCTGGGAACACAAAAT
GTTCTAAATGTCCTCCACACAGTTTAACATACATGGAAGCAACTTCTGTCTGTCAGTGTGAAAAGGGT
TATTTCCGAGCTGAAAAAGACCCACCTTCTATGGCATGTACCAGGCCACCTTCAGCTCCTAGGAATGT
GGTTTTTAACATCAATGAAACAGCCCTTATTTTGGAATGGAGCCCACCAAGTGACACAGGAGGGAGAA
AAGATCTCACATACAGTGTAATCTGTAAGAAATGTGGCTTAGACACCAGCCAGTGTGAGGACTGTGGT
GGAGGACTCCGCTTCATCCCAAGACATACAGGCCTGATCAACAATTCCGTGATAGTACTTGACTTTGT
GTCTCACGTGAATTACACCTTTGAAATAGAAGCAATGAATGGAGTTTCTGAGTTGAGTTTTTCTCCCA
AGCCATTCACAGCTATTACAGTGACCACGGATCAAGATGCACCTTCCCTGATAGGTGTGGTAAGGAAG
GACTGGGCATCCCAAAATAGCATTGCCCTATCATGGCAAGCACCTGCTTTTTCCAATGGAGCCATTCT
GGACTACGAGATCAAGTACTATGAGAAAGAACATGAGCAGCTGACCTACTCTTCCACAAGGTCCAAAG

SGK251 (SEQ ID NO: 9) (CONTINUED FROM PREVIOUS PAGE)

AGAGAGATCCCAGTTGCCATTAAAACTTTGAAAGGTGGCCACATGGATCGGCAAAGAAGAGATTTTCT
AAGAGAAGCTAGTATCATGGGCCAGTTTGACCATCCAAACATCATTCGCCTAGAAGGGGTTGTCACCA
AAAGATCCTTCCCGGCCATTGGGGTGGAGGCGTTTTGCCCCAGCTTCCTGAGGGCAGGGTTTTTAAAT
AGCATCCAGGCCCCGCATCCAGTGCCAGGGGGAGGATCTTTGCCCCCCAGGATTCCTGCTGGCAGACC
AGTAATGATTGTGGTGGAATATATGGAGAATGGATCCCTAGACTCCTTTTTGCGGAAGCATGATGGCC
ACTTCACAGTCATCCAGTTGGTCGGAATGCTCCGAGGCATTGCATCAGGCATGAAGTATCTTTCTGAT
ATGGGTTATGTTCATCGAGACCTAGCGGCTCGGAATATACTGGTCAATAGCAACTTAGTATGCAAAGT
TTCTGATTTTGGTCTCTCCAGAGTGCTGGAAGATGATCCAGAAGCTGCTTATACAACAACTGGTGGAA
AAATCCCCATAAGGTGGACAGCCCCAGAAGCCATCGCCTACAGAAAATTCTCCTCAGCAAGCGATGCA
TGGAGCTATGGCATTGTCATGTGGGAGGTCATGTCCTATGGAGAGAGACCTTATTGGGAAATGTCTAA
CCAAGATGTCATTCTGTCCATTGAAGAAGGGTACAGACTTCCAGCTCCCATGGGCTGTCCAGCATCTC
TACACCAGCTGATGCTCCACTGCTGGCAGAAGGAGAGAAATCACAGACCAAAATTTACTGACATTGTC
AGCTTCCTTGACAAACTGATCCGAAATCCCAGTGCCCTTCACACCCTGGTGGAGGACATCCTTGTAAT
GCCAGAGTCCCCTGGTGAAGTTCCGGAATATCCTTTGTTTGTCACAGTTGGTGACTGGCTAGATTCTA
TAAAGATGGGGCAATACAAGAATAACTTCGTGGCAGCAGGGTTTACAACATTTGACCTGATTTCAAGA
ATGAGCATTGAGGACATTAGAAGAATTGGAGTCATACTTATTGGACACCAGAGACGAATAGTCAGCAG
CATACAGACTTTACGTTTACACATGATGCACATACAGGAGAAGGGATTTCATGTATGAAAGTACCACA
AGCACCTGTGTTTTGTGCCTCAGCATTTCTAAAATGAACGATATCCTCTCTACTACTCTCTCTTCTGA
TTCTCCAAACATCACTTCACAAACTGCAGTCTTCTGTTCAGACTATAGGCACACACCTTATGTTTATG
CTTCCAACCAGGATTTTAAAATCATGCTACATAAATCCGTTCTGAATAACCTGCAACTAAAACCCTGG
CCCACTGCAGATTATTGCTACGCAATGCAACAGCTTTAAAACCTATCTCAGCCCATGAATGGAAAACA
AATCCAAATCCGATCCTTGAAAAGCAAAGGCTCTAAAGAAGCTCTTCAGAAGAGACGGTAAAGAATGA
ATTCTTTTACTTATCACCCAACCACATTTCTTAAAAATGTGTTTTGGTGTCTTTTCCTACCAAATTTC
TGCTCTACAAGGCAGTCAGTTAAATCTCTCATTTCATAATTTTCACTGTGCTAGTTCTTTCTTTTCTC
CTCTTTTAATAAATTTTCTAAAACTTT

SGK307 (SEQ ID NO: 10)

GTCGACCCACGCGTCCGCCCCGCACCCACCCACGTCTTCCCGGGAGTCGTATCCCGAGCATGGAGGTT
ACTGAGACCGTTATTTCTTCATGGCCTGCCTAGCTTAAGCAGTAGCTGGAAAAGATGTCTCGGGCTGT
TCGTCTTCCAGTCCCCTGTCCTGTTCAACTTGGTACCTTAAGAAATGACTCCCTGGAAGCTCAGCTTC
ATGAGTATGTCAAACAAGGGAACTATGTGAAAGTGAAGAAAATTCTTAAGAAAGGAATTTATGTTGAT
GCAGTTAACTCCTTGGGCCAAACAGCACTTTTTGTTGCGGCGTTATTGGGCCTTAGGAAATTCGTTGA
TGTTCTGGTGGATTATGGATCAGATCCAAATCACCGCTGCTTTGATGGGAGCACCCCTGTCCATGCAG
CAGCATTTTCGGGCAATCAGTGGATCCTTAGCAAACTGCTGGATGCAGGAGGTGACCTGCGACTCCAC
GATGAGAGGGGTCAAAACCCGAAGACTTGGGCTTTGACAGCAGGAAAGGAGCGTAGCACCCAGATAGT
GGAGTTCATGCAGCGCTGTGCCTCACACATGCAGGCCATCATCCAGGGCTTCTCTTACGACCTCCTGA
AGAAGATAGACTCCCCGCAGCGGCTTGTCTACAGCCCGTCCTGGTGTGGGGGCCTCGTGCAGGGAAAC
CCTAATGGCTCTCCTAACCGACTGCTTAAAGCTGGAGTCATTTCTGCTCAAAATATCTACAGCTTTGG
TTTTGGGAAGTTTTATCTTACTGGGGCGACACAGATGGCCTATCTAGGATCTCTTCCGGTCATTGGAG
AAAAGGAAGTGATTCAAGCTGATGATGAGCCCACCTTCTCTTTCTTCAGCGGCCCCTACATGGTCATG
ACCAACCTAGTGTGGAATGGGAGCAGGGTCACAGTGAAAGAGCTGAATCTCCCCACCCACCCACACTG
CAGCAGGCTGCGGCTGGCCGACTTGTTAATTGCCGAGCAGGAACACAGCAGCAAGCTGCGGCACCCCT
ACTTGCTACAGTTGATGGCTGTGTGTCTCTCCCAGGACCTAGAGAAAACCCGCCTTGTGTACGAGCGC
ATCACTATCGGCACATTGTTCAGTGTCCTTCATGAACGACGGTCCCAGTTCCCAGTGCTGCACATGGA
GGTGATTGTGCACCTGCTGCTCCAGATATCTGATGCCCTGAGATACCTGCATTTCCAGGGGTTTATCC
ACCGCTCCCTCAGCTCCTATGCTGTCCATATCATCTCCCCAGGTGAAGCGAGGCTGACCAACCTGGAG
TACATGTTGGAAAGCGAGGACAGAGGTGTACAGAGGGACCTGACTCGAGTGCCCCTTCCTACGCAGCT
ATACAACTGGGCCGCACCAGAAGTGATCTTACAGAAGGCAGCCACAGTGAAATCAGACATCTACAGCT
TTTCTATGATCATGCAGGAGATTTTAACAGATGACATACCCTGGAAGGGCTTAGATGGCTCAGTTGTT
AAAAAAGCCGTAGTCTCGGGGAATTATTTAGAAGCTGATGTCAGGCTTCCGAAACCTTACTATGATAT
TGTTAAGTCAGGCATCCACGTCAAGCAGAAAGACCGAACTATGAACCTTCAAGATATCCGGTATATTC
TGAAGAATGACTTAAAGGATTTTACTGGAGCCCAGAGAACTCAACCAACCGAGAGCCCCAGAGTGCAG
AGATACGGACTCCATCCCGATGTCAATGTCTATCTAGGACTGACTTCAGAACACCCCAGAGAGACACC
TGACATGGAAATCATAGAACTAAAGGAAATGGGCAGTCAACCTCATTCACCAAGGGTTCACTCTTTAT

Figure 1I

SGK307 (SEQ ID NO: 10) (CONTINUED FROM PREVIOUS PAGE)
TCACTGAGGGGACACTAGATCCTCAGGCCCCAGATCCATGTCTGATGGCCAGGGAGACTCAGAATCAA
GATGCTCCTTGCCCTGCTCCATTTATGGCAGAAGAGGCCAGCAGCCCCAGCACAGGTCAGCCAAGCCT
CTGCAGTTTCGAAATCAACGAGATCTACTCAGGCTGCTTGATTTTGGAAGATGACATAGAAGAGCCTC
CAGGAGCTGCTTCATCTTTGGAGGCAGACGGACCTAACCAGGTAGATGAACTGAAATCCATGGAAGAA
GAGCTGGATAAGATGGAGAGAGAGGCGTGTTGTTTTGGCAGTGAGGATGAGAGCTCTTCAAAAGCTGA
GACAGAGTACTCTTTTGATGACTGGGACTGGCAAAACGGTTCACTCAGTTCACTCAGCCTTCCTGAGT
CAACCAGAGAAGCCAAGAGCAATTTGAACAACATGTCCACGACTGAGGAGTATCTCATCAGTAAGTGT
GTGCTGGATCTAAAGATTATGCAGACAATAATGCACGAGAATGATGATAGGCTGAGGAATATCGAGCA
GATATTAGATGAAGTCGAGATGAAACAGAAGGAACAGGAAGAGCGCATGTCTTTATGGGCCACTTCAA
GAGAGTTTACAAATGCCTACAAGTTACCTCTGGCCGTGGGCCCTCCATCTTTAAACTATATTCCTCCT
GTCCTACAGCTTTCAGGGGGTCAGAAGCCAGACACCAGTGGCAACTACCCAACCCTACCAAGATTTCC
AAGAATGCTGCCGACTCTTTGTGACCCTGGAAAACAGAACACAGATGAACAATTTCAGTGCACTCAAG
GAGCCAAGGACAGTTTGGAAACAAGCAGGATCCAAAATACCAGTAGCCAGGGAAGACCTAGAGAGTCC
ACTGCCCAAGCCAAAGCCACACAGTTTAATAGTGCACTCTTCACTCTGTCAAGCCACCGGCAGGGACC
TTCTGCATCACCCAGCTGTCACTGGGACTCTACCAGGATGAGTGTGGAACCTGTTTCTTCTGAAATCT
ATAATGCAGAGTCCAGAAATAAAGATGATGGAAAGGTACACTTAAAATGGAAAATGGAGGTGAAAGAA
ATGGCAAAGAAAGCAGCTACTGGACAGCTCACAGTACCTCCTTGGCATCCTCAGAGTAGTCTGACTTT
AGAGAGCGAGGCTGAAAATGAGCCCGACGCCCTGCTGCAGCCCCCCATTAGGAGCCCAGAAAACACGG
ATTGGCAGCGAGTTATTGAGTATCATAGGGAAAATGATGAGCCCAGAGGAAATGGCAAGTTTGACAAG
ACGGGCAACAATGACTGTGACAGTGACCAGCATGGCAGACAGCCCAGGCTTGGAAGCTTCACCAGTAT
CAGGCACCCATCTCCCAGACAAAAGGAGCAACCAGAGCATAGTGAAGCCTTCCAAGCAAGTTCTGACA
CATTGGTGGCTGTAGAGAAATCTTACAGTACCTCGAGTCCCATAGAAGAGGACTTTGAAGGAATACAA
GGTGCATTTGCCCAACCTCAAGTCTCTGGTGAGGAAAAGTTCCAAATGAGAAAAATTCTTGGAAAGAA
TGCTGAGATTTTGCCCAGGTCTCAATTTCAACCTGTACGAAGTACTGAAGATGAACAAGAAGAGACAT
CAAAGGAGTCACCAAAGGAACTGAAAGAGAAAGACATATCATTGACGGATATTCAAGACCTGTCTAGT
ATCTCCTATGAACCAGACAGCTCTTTTAAGGAAGCTTCATGCAAAACACCCAAAATAAACCATGCACC
TACCAGTGTCAGCACTCCACTCAGCCCAGGGTCCGTTTCTTCAGCTGCCAGTCAGTATAAAGACTGCC
TTGAAAGTATCACATTTCAGGTTAAGACAGAGTTTGCCTCTTGCTGGAACAGTCAAGAATTTATTCAA
ACTTTGTCTGATGACTTTATAAGTGTCCGAGAGAGAGCAAAGAAACTGGATTCTCTCCTTACTTCCTC
TGAAACTCCCCCTTCAAGACTGACTGGTCTTAAAAGATTGTCTTCATTTATTGGGGCTGGATCCCCCA
GCCTTGTTAAGGCATGTGACTCATCACCACCCCATGCCACCCAGAGAAGGAGCCTGCCTAAAGAGCTC
TTGAAACAGCAGCAGGGCTCATCCACGGTGTTGCATGAGAACACAGCAAGTGATGGAGGAGGCACTGC
AAATGATCAAAGGCACTTAGAAGAACAAGAAACTGACAGTAAAAAAGAAGATAGTAGTATGCTTTTGT
CCAAAGAAACTGAAGATCTTGGAGAGGACACAGAGAGAGCTCACTCTACTCTGGATGAGGACCTGGAA
AGATGGCTGCAGCCACCTGAGGAGAGCGTGGAGCTACAAGACCTTCCCAAGGGCTCTGAAAGGGAGAC
AAATATCAAGGATCAAAAAGTTGGTGAAGAGAAAAGAAAAAGGGAAGATAGCATTACACCAGAGAGAA
GGAAATCAGAGGGTGTTCTAGGGACTTCTGAAGAAGATGAACTAAAATCCTGTTTTTGGAAGCGACTA
GGTTGGTCCGAATCATCCAGGATAATCGTGCTGGATCAGAGTGACTTGTCAGACTGATTGGAATTGGA
TCATAGACGGACTCCTGGCCTGAGTTTGAGTGTCCTGGTTGTAAGCTCCTTTCTTCTCTTTCTGCTTC
AGTTGCTGTCAGGGCAGCAGTTCCAGTTCTGTAAGTCTCACTTTGTTCAGCTGCCACAATAGACATCA
TCGTTTGGCCCTCTCTGTTAGCAGCACATTCAACCATTTGTTTTCAGTCAGATTTCTGAAAAGTGAGA
GGTAGTTTTGATAGTAAAAATTTTTGGTTGTGCCTAGAATGGCTTTGGTTTTGTTGATGTTAATTTTC
AAAAACTTTAACTCTTGTTATATAATAAAATGTTTAATTTTAATAAC

SGK119 (SEQ ID NO: 11)
CTGATGAAGAAGTGCTGGGACGAGGCTCCAGAGGACAGACCTAGCATGGACCAGATCTACAGCCAGTT
CAAAAGCATCAACCAAGGCAAGAGGACCAGTGTTGCTGACTCCATGCTGTGGTTGCTGGAGAAATATT
CCCAGAACCTGGAGGACCTGATTCAGGAGCAGACTGAGGAACTGGAGCTGAAGAGAGAGAAGACAGAA
AGGCTGCTCTGTCAGATGATTCCCCCGTCTGTGGCTGAAGCTCGGAAAATGGGGGCAACTGTGGAACC
AGGGTATTTTGACCAGGTTACCATATACTTCAGTGACATTGTGGGTTTCACCATCATCTCAGCCCTGA
GTGAACCCATTGAGGCGGTGGGCTTGCTCAACGATCTCTACATGCTGTTTGATGCTGTTCTGGGCAGC
CATGACCTGTATAAGGTGGAGACCATTGGGGACGCCTACATGGTGGTGTGGGGGCTGCCTCAGTGCAA
TGGCAGTCAGCATGCGGCCGAGATCAATAACATGGCTCTGGATATCCTTGGCTCTGTGGGTGACTTCC
GGAAGAGGCATGCACCCAACGTGCCCATTTGCATCAGGGCTGGCCTGCATTCAGGGCCCTGTGCAGCA

Figure 1J

SGK119 (SEQ ID NO: 11) (CONTINUED FROM PREVIOUS PAGE)
GGGGTCGTGGGTCTCACGATGCCTTGGTATTGCCTCTCTGGGGACACTGTCAACACGGCTTCCCAGAT
GGAGTCCACAGGGCTGCATGAAATTCATGTCAGCCAAAGCACTGTCCAGATACTGCTCAGCCTTGATG
AAAGCTACCAAATTGATATCAGAGGTCAGACTGAGCTGAAGGGGAAGGGCATCGAAGAGACCTACTGG
GTG

SGK387 (SEQ ID NO: 12)
ATGGTTGATGTTATCTGGCGTCAGTTGATATCAAGCTGCCCATGGCTTTCAGAACTTGATGAAAGTGC
AACTGAAGGAGTTATTAAAGTGTGGAGGAAAGTTGTAGATAGAATATTCAGCCTGTACAAACTCTCTT
GCAGTGCATACTTTACTTTCCTTAAACTCAACGCTGGTCAAATTCCTTTAGATGAGGATGACCCTAGG
CTGCATTTAAGTCACAGAGTGGAACAGAGCACTGATGACATGATTGTGATGGCCACATTGCGCCTGCT
GCGGTTGCTCGTGAAGCACGCTGGTGAGCTTCGGCAGTATCTGGAGCACGGCTTGGAGACAACACCCA
CTGCACCATGGAGAGGAATTATTCCGCAACTTTTCTCACGCTTAAACCACCCTGAAGTGTATGTGCGC
CAAAGTATTTGTAACCTTCTCTGCCGTGTGGCTCAAGATTCCCCACATCTCATATTGTATCCTGCAAT
AGTGGGAAGTCCTCCTGCATCTCAGGATAGCAATAAGGATGAACCTAAAAGTGGATTAAATGAAGACC
AAGCCATGATGCAGGATTGTTACAGCAAAATTGTAGATAAGCTGTCCTCTGCAAACCCCACCATGGTA
TTACAGGTTCAGATGCTCGTGGCTGAACTGCGCAGGGTCACTGTGCTCTGGGATGAGCTCTGGCTGGG
AGTTTTGCTGCAACAACACATGTATGTCCTGAGACGAATTCAGCAGCTTGAAGATGAGGTGAAGAGAG
TTCTTAACAACAACACCTTACGCAAAGAAGAGAAAATTGCAATCATGAGGGAGAAGCACACAGCTTTG
ATGAAGCCCATCGTATTTGCTTTGGAGCATGTGAGGAGTATCACAGCGGCTCCTGCAGAAACACCTCA
TGAAAAATGGTTTCAGGATAACTATGGTGATGCCATTGAAAATGCCCTAGAAAAACTGAAGACTCCAT
TGAACCCTGCAAAGCCTGGGAGCAGCTGGATTCCATTTAAAGAGATAATGCTAAGTTTGCAACAGAGA
GCACAGAAGCGTGCAAGTTACATCTTGCGTCTTGAAGAAATCAGTCCATGGTTGGCTGCCATGACTAA
CACTGAAATTGCTCTTCCTGGGGAAGTCTCAGCCAGAGACACTGTCACAATCCATAGTGTGGGCGGAA
CCATCACAATCTTACCGACTAAAACCAAGCCAAAGAAACTTCTCTTTCTTGGATCAGATGGGAAGAGC
TATCCTTATCTTTTCAAAGGACTGGAGGATTTACATCTGGATGAGAGAATAATGCAGTTCCTATCTAT
TGTGAATACCATGTTTGCTACAATTAATCGCCAAGAAACACCCCGGTTCCATGCTCGACACTATTCTG
TAACACCACTAGGAACAAGATCAGGACTAATCCAGTGGGTAGATGGAGCCACACCCTTATTTGGTCTT
TACAAACGATGGCAACAACGGGAAGCTGCCTTACAAGCACAAAAGAATCCTGGAATTGTACCCCGTCC
TAGTGAACTTTATTACAGTAAAATTGGCCCTGCTTTGAAAACAGTTGGGCTTAGCCTGGATGTGTCCC
GTCGGGATTGGCCTCTTCATGTAATGAAGGCAGTATTGGAAGAGTTAATGGAGGCCACACCCCCGAAT
CTCCTTGCCAAAGAGCTCTGGTCATCTTGCACAACACCTGATGAATGGTGGAGAGTTACGCAGTCTTA
TGCAAGATCTACTGCAGTCATGTCTATGGTTGGATACATAATTGGCCTTGGAGACAGACATCTGGATA
ATGTTCTTATAGATATGACGACTGGAGAAGTTGTTCACATAGATTACAATGTTTGCTTTGAAAAAGGT
AAAAGCCTTAGAGTTCCTGAGAAAGTACCTTTTCGAATGACACAAAACATTGAAACAGCACTGGGTGT
AACTGGAGTAGAAGGTGTATTTAGGCTTTCATGTGAGCAGGTTTTACACATTATGCGGCGTGGCAGAG
AGACCCTGCTGACGCTGCTGGAGGCCTTTGTGTACGACCCTCTGGTGGACTGGACAGCAGGAGGCGAG
GCTGGGTTTGCTGGTGCTGTCTATGGTGGAGGTGGCCAGCAGGCCGAGAGCAAGCAGAGCAAGAGAGA
GATGGAGCGAGAGATCACCCGCAGCCTGTTTTCTTCTAGAGTAGCTGAGATTAAGGTGAACTGGTTTA
AGAATAGAGATGAGATGCTGGTTGTGCTTCCCAAGTTGGACGGTAGCTTAGATGAATACCTAAGCTTG
CAAGAGCAACTGACAGATGTGGAAAAACTGCAGGGCAAACTACTGGAGGAAATAGAGTTTCTAGAAGG
AGCTGAAGGGGTGGATCATCCTTCTCATACTCTGCAACACAGGTATTCTGAGCACACCCAACTACAGA
CTCAGCAAAGAGCTGTTCAGGAAGCAATCCAGGTGAAGCTGAATGAATTTGAACAATGGATAACACAT
TATCAGGCTGCATTCAATAATTTAGAAGCAACACAGCTTGCAAGCTTGCTTCAAGAGATAAGCACACA
AATGGACCTTGGTCCTCCAAGTTACGTGCCAGCAACAGCCTTTCTGCAGAATGCTGGTCAGGCCCACT
TGATTAGCCAGTGCGAGCAGCTGGAGGGGGAGGTTGGTGCTCTCCTGCAGCAGAGGCGCTCCGTGCTC
CGTGGCTGTCTGGAGCAACTGCATCACTATGCAACCGTGGCCCTGCAGTATCCGAAGGCCATATTTCA
GAAACATCGAATTGAACAGTGGAAGACCTGGATGGAAGAGCTCATCTGTAACACCACAGTAGAGCGTT
GTCAAGAGCTCTATAGGAAATATGAAATGCAATATGCTCCCCAGCCACCCCCAACAGTGTGTCAGTTC
ATCACTGCCACTGAAATGACCCTGCAGCGATACGCAGCAGACATCAACAGCAGACTTATTAGACAAGT
GGAACGCTTGAAACAGGAAGCTGTCACTGTGCCAGTTTGTGAAGATCAGTTGAAAGAAATTGAACGTT
GCATTAAAGTTTTCCTTCATGAGAATGGAGAAGAAGGATCTTTGAGTCTAGCAAGTGTTATTATTTCT
GCCCTTTGTACCCTTACAAGGCGTAACCTGATGATGGAAGGTGCAGCGTCAAGTGCTGGAGAACAGCT
GGTTGATCTGACTTCTCGGGATGGAGCCTGGTTCTTGGAGGAACTCTGCAGTATGAGCGGAAACGTCA
CCTGCTTGGTTCAGTTACTGAAGCAGTGCCACCTGGTGCCACAGGACTTAGATATCCCGAACCCCATG

Figure 1K

SGK387 (SEQ ID NO: 12) (CONTINUED FROM PREVIOUS PAGE)
GAAGCGTCTGAGACAGTTCACTTAGCCAATGGAGTGTATACCTCACTTCAGGAATTGAATTCGAATTT
CCGGCAAATCATATTTCCAGAAGCACTTCGATGTTTAATGAAAGGGGAATACACGTTAGAAAGTATGC
TGCATGAACTGGACGGTCTTATTGAGCAGACCACCGATGGCGTTCCCCTGCAGACTCTAGTGGAATCT
CTTCAGGCCTACTTAAGAAACGCAGCTATGGGACTGGAAGAAGAAACACATGCTCATTACATCGATGT
TGCCAGACTACTACATGCTCAGTACGGTGAATTAATCCAACCGAGAAATGGTTCAGTTGATGAAACAC
CCAAAATGTCAGCTGGCCAGATGCTTTTGGTAGCATTCGATGGCATGTTTGCTCAAGTTGAAACTGCT
TTCAGCTTATTAGTTGAAAAGTTGAACAAGATGGAAATTCCCATAGCTTGGCGAAAGATTGACATCAT
AAGGGAAGCCAGGAGTACTCAAGTTAATTTTTTTGATGATGATAATCACCGGCAGGTGCTAGAAGAGA
TTTTCTTTCTAAAAAGACTACAGACTATTAAGGAGTTCTTCAGGCTCTGTGGTACCTTTTCTAAAACA
TTGTCAGGATCAAGTTCACTTGAAGATCAGAATACTGTGAATGGGCCTGTACAGATTGTCAATGTGAA
AACCCTTTTTAGAAACTCTTGTTTCAGTGAAGACCAAATGGCCAAACCTATCAAGGCATTCACAGCTG
ACTTTGTGAGGCAGCTCTTGATAGGGCTACCCAACCAAGCCCTCGGACTCACACTGTGCAGTTTTATC
AGTGCTCTGGGTGTAGACATCATTGCTCAAGTAGAGGCAAAGGACTTTGGTGCCGAAAGCAAAGTTTC
TGTTGATGATCTCTGTAAGAAAGCGGTGGAACATAACATCCAGATAGGGAAGTTCTCTCAGCTGGTTA
TGAACAGGGCAACTGTGTTAGCAAGTTCTTACGACACTGCCTGGAAGAAGCATGACTTGGTGCGAAGG
CTAGAAACCAGTATTTCTTCTTGTAAGACAAGCCTGCAGCGGGTTCAGCTGCATATTGCCATGTTTCA
GTGGCAACATGAAGATCTACTTATCAATAGACCACAAGCCATGTCAGTCACACCTCCCCCACGGTCTG
CTATCCTAACCAGCATGAAAAAGAAGCTGCATACCCTGAGCCAGATTGAAACTTCTATTGCGACAGTT
CAGGAGAAGCTAGCTGCACTTGAATCAAGTATTGAACAGCGACTCAAGTGGGCAGGTGGTGCCAACCC
TGCATTGGCCCCTGTACTACAAGATTTTGAAGCAACGATAGCTGAAAGAAGAAATCTTGTCCTTAAAG
AGAGCCAAAGAGCAAGTCAGGTCACATTTCTCTGCAGCAATATCATTCATTTTGAAAGTTTACGAACA
AGAACTGCAGAAGCCTTAAACCTGGATGCGGCGTTATTTGAACTAATCAAGCGATGTCAGCAGATGTG
TTCGTTTGCATCACAGTTTAACAGTTCAGTGTCTGAGTTAGAGCTTCGTTTATTACAGAGAGTGGACA
CTGGTCTTGAACATCCTATTGGCAGCTCTGAATGGCTTTTGTCAGCACACAAACAGTTGACCCAGGAT
ATGTCTACTCAGAGGGCAATTCAGACAGAGAAAGAGCAGCAGATAGAAACGGTCTGTGAAACAATTCA
GAATCTGGTTGATAATATAAAGACTGTGCTCACTGGTCATAACCGACAGCTTGGAGATGTCAAACATC
TCTTGAAAGCTATGGCTAAGGATGAAGAAGCTGCTCTGGCAGATGGTGAAGATGTTCCCTATGAGAAC
AGTGTTAGGCAGTTTTTGGGTGAATATAAATCATGGCAAGACAACATTCAAACAGTTCTATTTACATT
AGTCCAGGCTATGGGTCAGGTTCGAAGTCAAGAACACGTTGAAATGCTCCAGGAAATCACTCCCACCT
TGAAAGAACTGAAAACACAAAGTCAGAGTATCTATAATAATTTAGTGAGTTTTGCATCACCCTTAGTC
ACCGATGCAACAAATGAATGTTCGAGTCCAACGTCATCTGCTACTTATCAGCCATCCTTCGCTGCAGC
AGTCCGGAGTAACACTGGCCAGAAGACTCAGCCTGATGTCATGTCACAGAATGCTAGAAAGCTGATCC
AGAAAAATCTTGCTACATCAGCTGATACTCCACCAAGCACCGTTCCAGGAACTGGCAAGAGTGTTGCT
TGTAGTCCTAAAAAGGCAGTCAGAGACCCTAAAACTGGGAAAGCGGTGCAAGAGAGAAACTCCTATGC
AGTGAGTGTGTGGAAGAGAGTGAAAGCCAAGTTAGAGGGCCGAGATGTTGATCCGAATAGGAGGATGT
CAGTTGCTGAACAGGTTGACTATGTCATTAAGGAAGCAACTAATCTAGATAACTTGGCTCAGCTGTAT
GAAGGTTGGACAGCCTGGGTGTGA

Figure 1L

SGK216 (SEQ ID NO: 13)
AMIVRNAKDTAHTKAEWNILEEVRHPFIVDLTYAFQTGGKLYLILEYLSGGELFMQLEREGIFMEDTA
CFYLAEISMALGHLHQKGIIYRDLKPENIILNHQGHMKLTDFGLRKESIHDGTVTHTFCGTIEYMAPE
ILMRSGHNRAVDWWSLGALMYDMLTGAPPFTGENRKKTIDNILKCKLNLPPYLTQEARDLLKKLLKRN
AASLGAGPGDAGEVQAHPFFRHINWEELLAQKVEPPFKPLLQSEEDVSQFDSKFTRQTPVDSPDDATL
SETANQVFL

SGK237 (SEQ ID NO: 14)
MRVLFDESVLPPTVYFKNCSILFLASLCAFGVLTGLLVWSFMQYMEIVANEYLGYGEEQHTVDKLVNM
TYIFQKLAAVKDQREWVTTSGAHKTLVNLLGARDTNVLLGSLLALASLAERLTAELLRLLCAEPQVKE
QVKLYEGIPVLLSLLHSDHLKLLWSIVWILVQVCEDPETSVEIRIWGGIKQLLHILQGDRNFVSDHSS
IGSLSSANAAGRIQQLHLSEDLSPREIQENTFSLQAACCAALTELVLNDTNAHQVVQENGVYTIAKLI
LPNKQKNAAKSNLLQCYAFRALRFLFSMERNRPLFKRLFPTDLFEIFIDIGHYVRDISAYEELVSKLN
LLVEDELKQIAENIESINQNKAPLKYIGNYAILDHLGSGAFGCVYKVRKHSGQNLLAMKEVNLHNPAF
GKDKKDRDSSVRNIVSELTIIKEQLYHPNIVRYYKTFLENDRLYIVMELIEGAPLGEHFSSLKEKHHH
FTEERLWKIFIQLCLALRYLHKEKRIVHRDQTPNNIMLGDKDKVTVTDFGLAKQKQENSKLTSVVGTI
LYSCPEVLKSEPYGEKADVWAVGCILYQMATLSPPFYSTNMLSLATKIVEAVYEPVPEGIYSEKVTDT
ISRCLTPDAEARPDIVEVSSMISDVMMKYLDNLSTSQLSLEKKLERERRRTQRYFMEANRNTVTCHHE
LAVLSHETFEKASLSSSSSGAASLKSELSESADLPPEGFQASYGKDEDRACNEILSDDNFNLENAEKD
TYSEVDDELDISDNSSSSSSSPLKESTFNILKRSFSASGGERQSQTRDFTGGTGSRPRPGPQMGTFLW
QASAGIAVSQRKVRQISDPIQQILIQLHKIIYITQLPP

SGK248 (SEQ ID NO: 15)
MDHPSREKDERQRTTKPMAQRSAHCSRPSGSSSSSGVLMVGPNFRVGKKIGCGNFGELRLGKNLYTNE
YVAIKLEPIKSRAPQLHLEYRFYKQLGSAGEGLPQVYYFGPCGKYNAMVLELLGPSLEDLFDLCDRTF
TLKTVLMIAIQLLSRMEYVHSKNLIYRDVKPENFLIGRQGNKKEHVIHIIDFGLAKEYIDPETKKHIP
YREHKSLTGTARYMSINTHLGKEQSRRDDLEALGHMFMYFLRGSLPWQGLKADTLKERYQKIGDTKRN
TPIEALCENFPEEMATYLRYVRRLDFFEKPDYEYLRTLFTDLFEKKGYTFDYAYDWVGRPIPTPVGSV
HVDSGASAITRESHTHRDRPSQQQPLRNQVVSSTNGELNVDDPTGAHSNAPITAHAEVEVVEEAKCCC
FFKRKRKKTAQRHK

SGK223 (SEQ ID NO: 16)
MVGLHNLEPRGERNIAFHPVSFPEEKAVHKEKPSFPYQDRPSTQESFRQKLAAFAGTTSGCHQGPGPL
RESLPSEDDSDQRCSPSGDSEGGEYCSILDCCPGSPVAKAASQTAGSRGRHGGRDCSPTCWEQGKCSG
PAEQEKRGPSFPKECCSQGPTAHPSCLGPKKLSLTSEAAISSDGLSCGSGSGSGSGASSPFVPHLESD
YCSLMKEPAPEKQQDPGCPGVTPSRCLGLTGEPQPPAHPREATQPEPIYAESTKRKKAAPVPSKSQAK
IEHAAAAQGQGQVCTGNAWAQKAASGWGRDSPDPTPQVSATITVMAAHPEEDHRTIYLSSPDSAVGVQ
WPRGPVSQNSEVGEEETSAGQGLSSRESHAHSASESKPKERPAIPPKLSKSSPVGSPVSPSAGGPPVS
PLADLSDGSSGGSSIGPQPPSQGPADPAPSCRTNGVAISDPSRCPQPAASSASEQRRPRFQAGTWSRQ
CRIEEEEEVEQELLSHSWGRETKNGPTDHSNSTTWHRLHPTDGSSGQNSKVGTGMSKSASFAFEFPKD
RSGIETFSPPPPPPKSRHLLKMNKSSSDLEKVSQGSAESLSPSFRGVHVSFTTGSTDSLASDSRTCSD
GGPSSELAHSPTNSGKKLFAPVPFPSGSTEDVSPSGPQQPPPLPQKKIVSRAASSPDGFFWTQGSPKP
GTASPKLNLSHSETNVHDESHFSYSLSPGNRHHPVFSSSDPLEKAFKGSGHWLPAAGLAGNRGGCGSP
GLQCKGAPSASSSQLSVSSQASTGSTQLQLHGLLSNISSKEGTYAKLGGLYTQSLARLVAKCEDLFMG
GQKKELHFNENNWSLFKLTCNKPCCDSGDAIYYCATCSEDPGSTYAVKICKAPEPKTVSYCSPSVPVH
FNIQQDCGHFVASVPSSMLSSPDAPKDPVPALPTHPPAQEQDCVVVITREVPHQTASDFVRDSAASHQ
AEPEAYERRVCFLLLQLCNGLEHLKEHGIIHRDLCLENLLLVHCTLQAGPGPAPAPAPAPAAAAPPCS
SAAPPAGGTLSPAAGPASPEGPREKQLPRLIISNFLKAKQKPGGTPNLQQKKSQARLAPEIVSASQYR
KFDEFQTGILIYELLHQPNPFEVRAQLRERDYRQEDLPPLPALSLYSPGLQQLAHLLLEADPIKRIRI
GEAKRVLQCLLWGPRRELVQQPGTSEEALCGTLHNWIDMKRALMMMKFAEKAVDRRRGVELEDWLCCQ
YLASAEPGALLQSLKLLQLL

SGK269 (SEQ ID NO: 17)
MSACNTFTEHVWKPGECKNCFKPKSLHQLPPDPEKAPITHGNVKTNANHSNNHRIRNTGNFRPPVAKK
PTIAVKPTMIVADGQSICGELSIQEHCENKPVIIGWNRNRAALSQKPLNNNNEDDEGISHVPKPYGNN

Figure 2A

SGK269 (SEQ ID NO: 17) (CONTINUED FROM PREVIOUS PAGE)
DSAKKMSDNNNGLTEVLKEIAGLDTAPQIRGNETNSRETFLGRINDCYKRSLERKLPPSCMIGGIKET
QGKHVILSGSTEVISNEGGRFCYPEFSSGEESEEDVLFSNMEEEHESWDESDEELLAMEIRMRGQPRF
ANFRANTLSPVRFFVDKKWNTIPLRNKSLQRICAVDYDDSYDEILNGYEENSVVSYGQGSIQSMVSSD
STSPDSSLTEESRSETASSLSQKICNGGLSPGNPGDSKDMKEIEPNYESPSSNNQDKDSSQASKSSIK
VPETHKAVLALRLEEKDGKIAVQTEKEESKASTDVAGQAVTINLVPTEEQAKPYRVVNLEQPLCKPYT
VVDVSAAMASEHLEGPVNSPKTKSSSSTPNSPVTSSSLTPGQISAHFQKSSAIRYQEVWTSSTSPRQK
IPKVELITSGTGPNVPPRKNCHKSAPTSPTATNISSKTIPVKSPNLSEIKFNSYNNAGMPPFPIIIHD
EPTYARSSKNAIKVPIVINPNAYDNLAIYKSFLGTSGELSVKEKTTSVISHTYEEIETESKVPDNTTS
KTTDCLQTKGFSNSTEHKRGSVAQKVQEFNNCLNRGQSSPQRSYSSSHSSPAKIQRATQEPVAKIEGT
QESQMVGSSSTREKASTVLSQIVASIQPPQSPPETPQSGPKACSVEELYAIPPDADVAKSTPKSTPVR
PKSLFTSQPSGEAEAPQTTDSPTTKVQKDPSIKPVTPSPSKLVTSPQSEPPAPFPPPRSTSSPYHAGN
LLQRHFTNWTKPTSPTRSTEAESVLHSEGSRRAADAKPKRWISFKSFFRRRKTDEEDDKEKEREKGKL
VGLDGTVIHMLPPPPVQRHHWFTEAKGESSEKPAIVFMYRCDPAQGQLSVDQSKARTDQAAVMEKGRA
ENALLQDSEKKRSHSSPSQIPKKILSHMTHEVTEDFSPRDPRTVVGKQDGKGCTSVTTALSLPELERE
DGKEDISDPMDPNPCSATYSNLGQSRAAMIPPKQPRQPKGAVDDAIAFGGKTDQEAPNASQPTPPPLP
KKMIIRANTEPISKDLQKSMESSLCVMANPTYDIDPNWDASSAGSSISYELKGLDIESYDSLERPLRK
ERPVPSAANSISSLTTLSIKDRFSNSMESLSSRRGPSCRQGRGIQKPQRQALYRGLENREEVVGKIRS
LHTDALKKLAVKCEDLFMAGQKDQLRFGVDSWSDFRLTSDKPCCEAGDAVYYTASYAKDPLNNYAVKI
CKSKAKESQQYYHSLAVRQSLAVHFNIQQDCGHFLAEVPNRLLPWEDPDDPEKDEDDMEETEEDAKGE
TDGKNPKPCSEAASSQKENQGVMSKKQRSHVVVITREVPCLTVADFVRDSLAQHGKSPDLYERQVCLL
LLQLCSGLEHLKPYHVTHCDLRLENLLLVHYQPGGTAQGFGPAEPSPTSSYPTRLIVSNFSQAKQKSH
LVDPEILRDQSRLAPEIITATQYKKCDEFQTGILIYEMLHLPNPFDENPELKEREYTRADLPRIPFRS
PYSRGLQQLASCLLNPNPSERILISDAKGILQCLLWGPREDLFQTFTACPSLVQRNTLLQNWLDIKRT
LLMIKFAEKSLDREGGISLEDWLCAQYLAFATTDSLSCIVKILQHR

SGK258 (SEQ ID NO: 18)
METLNGAGDTGGKPSTRGGDPAARSRRTEGIRAAYRRGDRGGARDLLEEACDQCASQLEKGQLLSIPA
AYGDLEMVRYLLSKRLVELPTEPTDDNPAVVAAYFGHTAVVQNTLPTEPTDDNPAVVAAYFGHTAVVQ
ELLESLPGPCSPQRLLNWMLALACQRGHLGVVKLLVLTHGADPESYAVRKNEFPVIVRLPLYAAIKSG
NEDIAIFLLRHGAYFCSYILLDSPDPSKHLLRKYFIEASPLPSSYPGKTALRVKWSHLRLPWVDLDWL
IDISCQITELDLSANCLATLPSVIPWGLINLRKLNLSDNHLGELPGVQSSDEIICSRLLEIDISSNKL
SHLPPGFLHLSKLQKLTASKNCLEKLFEEENATNWIGLRKLQELDISDNKLTELPALFLHSFKSLNSL
NVSRNNLKVFPDPWACPLKCCKASRNALECLPDKMAVFWKNHLKDVDFSENALKEVPLGLFQLDALMF
LRLQGNQLAALPPQEKWTCRQLKTLDLSRNQLGKNEDGLKTKRIAFFTTRGRQRSGTEAETTMEFSAS
LVTIVFLSNNCNLCAYTCAASVLEFPAFLSESLEVLCLNDNHLDTVPPSVCLLKSLSELYLGNNPGLR
ELPPELGQLGNLWQLDTEDLTISNVPAEIQKEGPKAMLSYLRAQLRKAEKCKLMKMIIVGPPRQGKST
LLEILQTGRAPQVVHGEATIRTTKWELQRPAGSRAKVKDGLRAESLWVESVEFNVWDIGGPASMATVN
QCFFTDKALYVVVWNLALGEEAVANLQFWLLNIEAKAPNAVVLVVGTHLDLIEAKFRVERIATLRAYV
LALCRSPSGSRATGFPDITFKHLHEISCKSLEGQEGLRQLIFHVTCSMKDVGSTIGCQRLAGRLIPRS
YLSLQEAVLAEQQRRSRDDDVQYLTDRQLEQLVEQTPDNDIKDYEDLQSAISFLIETGTLLHFPDTSH
GLRNLYFLDPIWLSECLQRIFNIKGSRSVAKNGVIRAEDLRMLLVGTGFTQQTEEQYFQFLAKFEIAL
PVANDSYLLPHLLPSKPGLDTHGMRHPTANTIQRVFKMSFVPVGFWQRFIARMLISLAEMDLQLFENK
KNTKSRNRKVTIYSFTGNQRNRCSTFRVKRNQTIYWQEGLLVTFDGGYLSVESSDVNWKKKKSGGMKI
VCQSEVRDFSAMAFITDHVNSLIDQWFPALTATESDGTPLMEQYVPCPVCETAWAQHTDPSEKSEDVQ
YFDMEDCVLTAIERDFISCPRHPDLPVPLQELVPELFMTDFPARLFLENSKLEHSEDEGSVLGQGGSG
TVIYRARYQGQPVAVKRFHIKKFKNFANVPADTMLRHLRATDAMKNFSEFRQEASMLHALQHPCIVAL
IGISIHPLCFALELAPLSSLNTVLSENARDSSFIPLGHMLTQKIAYQIASGLAYLHKKNIIFCDLKSD
NILVWSLDVKEHINIKLSDYGISRQSFHEGALGVEGTPGYQAPEIRPRIVYDEKVDMFSYGMVLYELL
SGQRPALGHHQLQIAKKLSKGIRPVLGQPEEVQFRRLQALMMECWDTKPEKRPLALSVVSQMKDPTFA
TFMYELCCGKQTAFFSSQGQEYTVVFWDGKEESRNYTVVNTEKGLMEVQRMCCPGMKVSCQLQVQRSL
WTATENSYLVLAGLADGLVAVFPVVRGTPKDSCSYLCSHTANRSKFSIADEDARQNPYPVKAMEVVNS
GSEVWYSNGPGLLVIDCASLEICRRLEPYMAPSMVTSVVCSSEGRGEEVVWCLDDKANSLVMYHSTTY
QLCARYFCGVPSPLRDMFPVRPLDTEPPAASHTANPKVPEGDSIADVSIMYSEELGTQILIHQESLTD
YCSMSSYSSSPPRQAARSPSSLPSSPASSSSVPFSTDCEDSDMLHTPGAASDRSEHDLTPMDGETFSQ

Figure 2B

SGK258 (SEQ ID NO: 18) (CONTINUED FROM PREVIOUS PAGE)
HLQAVKILAVRDLIWVPRRGGDVIVIGLEKDSGAQRGRVIAVLKARELTPHGVLVDAAVVAKDTVVCT
FENENTEWCLAVWRGWGAREFDIFYQSYEELGRLEACTRKRR

SGK382 (SEQ ID NO: 19)
MFPLKDAEMGAFTFFASALPHDVCGSNGLPLTPNSIKILGRFQILKTITHPRLCQYVDISRGKHERLV
VVAEHCERSLEDLLRERKPVSCSTVLCIAFEVLQGLQYMNKHGIVHRALSPHNILLDRKGHIKLAKFG
LYHMTAHGDDVDFPIGYPSYLAPEVIAQGIFKTTDHMPSKKPLPSGPKSDVWSLGIILFELCVGRKLF
QSLDISERLKFLLTLDCVDDTLIVLAEEHGCLDIIKELPETVIDLLNKCLTFHPSKRPTPDELMKDKV
FSEVSPLYTPFTKPASLFSSSLRCADLTLPEDISQLCKDINNDYLAERSIEEVYYLWCLAGGDLEKEL
VNKEIIRSKPPICTLPNFLFEDGESFGQGRDRSSLLDDTTVTLSLCQLRNRLKDVGGEAFYPLLEDDQ
SNLPHSNSNNELSAAATLPLIIREKDTEYQLNRIILFDRLLKAYPYKKNQIWKEARVDIPPLMRGLTW
AALLGVEGAIHAKYDAIDKDTPIPTDRQIEVDIPRCHQYDELLSSPEGHAKFRRVLKAWVVSHPDLVY
WQGLDSLCAPFLYLNFNNEALAYACMSAFIPKYLYNFFLKDNSHVIQEYLTVFSQMIAFHDPELSNHL
NEIGFIPDLYAIPWFLTMFTHVFPLHKIFHLWDTLLLGNSSFPFCIGVAILQQLRDRLLANGFNECIL
LFSDLPEIDIERCVRESINLFCWTPKSATYRQHAQPPKPSSDSSGGRSSAPYFSAECPDPPKTDLSRE
SIPLNDLKSEVSPRISAEDLIDLCELTVTGHFKTPSKKTKSSKPKLLVVDIRNSEDFIRGHISGSINI
PFSAAFTAEGELTQGPYTAMLQNFKGKVIVIVGHVAKHTAEFAAHLVKMKYPRICILDGGINKIKPTG
LLTIPSPQI

SGK424 (SEQ ID NO: 20)
MEPTEPQPTDALAALSRFPVELGSKRDLDTQVGRLHHLAVAGGPGWGPARLLRRGEGLAAGVQVDGEN
SEGQTALFLSALLGHGSAMQLLLAFGAKPNQWLASHPFPNRLPPRCSALTAAWMAAPPCLQAPSLVMV
PLLQAGGDLQLHDQRGRTPQDWAEQGGAKQNWELLELLQLCRAHISAFVHDGELPPMASLDQLQARFG
NSPPGSLSSLRLMQADRMLRLAQIRRASQVPALGFGQLSSLWPPGLVTGIPLVDPKELLAAQGEPDCT
YESSSPPSWPRKAPSSQACTAPPLFSSLLWMGHKMTVQQPKAPGTQTDVLLADLQHCSKLHHPGLLLL
MALSPSADLSGLCLLFEPVWLGSLQGLVQPGLAKVGHLEHRRLLRQRWLRPRQQKGYPWGGPGPGLPP
PPELHPWLPLQLICGDMPTTTSDLYSFCILIQEVFTGELPWAGRGGPEVKAKLEAGESPALEPRVPAP
YQALIRAGLGLAPADRWGSLQSTRYLLREAPAQRNCTSSGSQSQDHTQACGPCDVPRPHPMPGPLEHH
GYQVAAPAFSPAGTRY

SGK251 (SEQ ID NO: 21)
MGGCEVREFLLQFGFFLPLLTAWPGDCSHVSNNQVVLLDTTTVLGELGWKTYPLNGWDAITEMDEHNR
PIHTYQVCNVMEPNQNNWLRTNWISRDAAQKIYVEMKFTLRDCNSIPWVLGTCKETFNLFYMESDESH
GIKFKPNQYTKIDTIAADESFTQMDLGDRILKLNTEIREVGPIERKGFYLAFQDIGACIALVSVRVFY
KKCPFTVRNLAMFPDTIPRVDSSSLVEVRGSCVKSAEERDTPKLYCGADGDWLVPLGRCICSTGYEEI
EGSCHACRPGFYKAFAGNTKCSKCPPHSLTYMEATSVCQCEKGYFRAEKDPPSMACTRPPSAPRNVVF
NINETALILEWSPPSDTGGRKDLTYSVICKKCGLDTSQCEDCGGGLRFIPRHTGLINNSVIVLDFVSH
VNYTFEIEAMNGVSELSFSPKPFTAITVTTDQDAPSLIGVVRKDWASQNSIALSWQAPAFSNGAILDY
EIKYYEKEHEQLTYSSTRSKAPSVIITGLKPATKYVFHIRVRTATGYSGYSQKFEFETGDETSDMAAE
QGQILVIATAAVGGFTLLVILTLFFLITGRCQWYIKAKMKSEEKRRNHLQNGHLRFPGIKTYIDPDTY
EDPSLAVHEFAKEIDPSRIRIERVIGAGEFGEVCSGRLKTPGKREIPVAIKTLKGGHMDRQRRDFLRE
ASIMGQFDHPNIIRLEGVVTKRSFPAIGVEAFCPSFLRAGFLNSIQAPHPVPGGGSLPPRIPAGRPVM
IVVEYMENGSLDSFLRKHDGHFTVIQLVGMLRGIASGMKYLSDMGYVHRDLAARNILVNSNLVCKVSD
FGLSRVLEDDPEAAYTTTGGKIPIRWTAPEAIAYRKFSSASDAWSYGIVMWEVMSYGERPYWEMSNQD
VILSIEEGYRLPAPMGCPASLHQLMLHCWQKERNHRPKFTDIVSFLDKLIRNPSALHTLVEDILVMPE
SPGEVPEYPLFVTVGDWLDSIKMGQYKNNFVAAGFTTFDLISRMSIEDIRRIGVILIGHQRRIVSSIQ
TLRLHMMHIQEKGFHV

SGK307 (SEQ ID NO: 22)
MSRAVRLPVPCPVQLGTLRNDSLEAQLHEYVKQGNYVKVKKILKKGIYVDAVNSLGQTALFVAALLGL
RKFVDVLVDYGSDPNHRCFDGSTPVHAAAFSGNQWILSKLLDAGGDLRLHDERGQNPKTWALTAGKER
STQIVEFMQRCASHMQAIIQGFSYDLLKKIDSPQRLVYSPSWCGGLVQGNPNGSPNRLLKAGVISAQN
IYSFGFGKFYLTGATQMAYLGSLPVIGEKEVIQADDEPTFSFFSGPYMVMTNLVWNGSRVTVKELNLP
THPHCSRLRLADLLIAEQEHSSKLRHPYLLQLMAVCLSQDLEKTRLVYERITIGTLFSVLHERRSQFP

Figure 2C

SGK307 (SEQ ID NO: 22) (CONTINUED FROM PREVIOUS PAGE)
VLHMEVIVHLLLQISDALRYLHFQGFIHRSLSSYAVHIISPGEARLTNLEYMLESEDRGVQRDLTRVP
LPTQLYNWAAPEVILQKAATVKSDIYSFSMIMQEILTDDIPWKGLDGSVVKKAVVSGNYLEADVRLPK
PYYDIVKSGIHVKQKDRTMNLQDIRYILKNDLKDFTGAQRTQPTESPRVQRYGLHPDVNVYLGLTSEH
PRETPDMEIIELKEMGSQPHSPRVHSLFTEGTLDPQAPDPCLMARETQNQDAPCPAPFMAEEASSPST
GQPSLCSFEINEIYSGCLILEDDIEEPPGAASSLEADGPNQVDELKSMEEELDKMEREACCFGSEDES
SSKAETEYSFDDWDWQNGSLSSLSLPESTREAKSNLNNMSTTEEYLISKCVLDLKIMQTIMHENDDRL
RNIEQILDEVEMKQKEQEERMSLWATSREFTNAYKLPLAVGPPSLNYIPPVLQLSGGQKPDTSGNYPT
LPRFPRMLPTLCDPGKQNTDEQFQCTQGAKDSLETSRIQNTSSQGRPRESTAQAKATQFNSALFTLSS
HRQGPSASPSCHWDSTRMSVEPVSSEIYNAESRNKDDGKVHLKWKMEVKEMAKKAATGQLTVPPWHPQ
SSLTLESEAENEPDALLQPPIRSPENTDWQRVIEYHRENDEPRGNGKFDKTGNNDCDSDQHGRQPRLG
SFTSIRHPSPRQKEQPEHSEAFQASSDTLVAVEKSYSTSSPIEEDFEGIQGAFAQPQVSGEEKFQMRK
ILGKNAEILPRSQFQPVRSTEDEQEETSKESPKELKEKDISLTDIQDLSSISYEPDSSFKEASCKTPK
INHAPTSVSTPLSPGSVSSAASQYKDCLESITFQVKTEFASCWNSQEFIQTLSDDFISVRERAKKLDS
LLTSSETPPSRLTGLKRLSSFIGAGSPSLVKACDSSPPHATQRRSLPKELLKQQQGSSTVLHENTASD
GGGTANDQRHLEEQETDSKKEDSSMLLSKETEDLGEDTERAHSTLDEDLERWLQPPEESVELQDLPKG
SERETNIKDQKVGEEKRKREDSITPERRKSEGVLGTSEEDELKSCFWKRLGWSESSRIIVLDQSDLSD

SGK119 (SEQ ID NO: 23)
LMKKCWDEAPEDRPSMDQIYSQFKSINQGKRTSVADSMLWLLEKYSQNLEDLIQEQTEELELKREKTE
RLLCQMIPPSVAEARKMGATVEPGYFDQVTIYFSDIVGFTIISALSEPIEAVGLLNDLYMLFDAVLGS
HDLYKVETIGDAYMVVWGLPQCNGSQHAAEINNMALDILGSVGDFRKRHAPNVPICIRAGLHSGPCAA
GVVGLTMPWYCLSGDTVNTASQMESTGLHEIHVSQSTVQILLSLDESYQIDIRGQTELKGKGIEETYW
V

SGK387 (SEQ ID NO: 24)
MVDVIWRQLISSCPWLSELDESATEGVIKVWRKVVDRIFSLYKLSCSAYFTFLKLNAGQIPLDEDDPR
LHLSHRVEQSTDDMIVMATLRLLRLLVKHAGELRQYLEHGLETTPTAPWRGIIPQLFSRLNHPEVYVR
QSICNLLCRVAQDSPHLILYPAIVGSPPASQDSNKDEPKSGLNEDQAMMQDCYSKIVDKLSSANPTMV
LQVQMLVAELRRVTVLWDELWLGVLLQQHMYVLRRIQQLEDEVKRVLNNNTLRKEEKIAIMREKHTAL
MKPIVFALEHVRSITAAPAETPHEKWFQDNYGDAIENALEKLKTPLNPAKPGSSWIPFKEIMLSLQQR
AQKRASYILRLEEISPWLAAMTNTEIALPGEVSARDTVTIHSVGGTITILPTKTKPKKLLFLGSDGKS
YPYLFKGLEDLHLDERIMQFLSIVNTMFATINRQETPRFHARHYSVTPLGTRSGLIQWVDGATPLFGL
YKRWQQREAALQAQKNPGIVPRPSELYYSKIGPALKTVGLSLDVSRRDWPLHVMKAVLEELMEATPPN
LLAKELWSSCTTPDEWWRVTQSYARSTAVMSMVGYIIGLGDRHLDNVLIDMTTGEVVHIDYNVCFEKG
KSLRVPEKVPFRMTQNIETALGVTGVEGVFRLSCEQVLHIMRRGRETLLTLLEAFVYDPLVDWTAGGE
AGFAGAVYGGGGQQAESKQSKREMEREITRSLFSSRVAEIKVNWFKNRDEMLVVLPKLDGSLDEYLSL
QEQLTDVEKLQGKLLEEIEFLEGAEGVDHPSHTLQHRYSEHTQLQTQQRAVQEAIQVKLNEFEQWITH
YQAAFNNLEATQLASLLQEISTQMDLGPPSYVPATAFLQNAGQAHLISQCEQLEGEVGALLQQRRSVL
RGCLEQLHHYATVALQYPKAIFQKHRIEQWKTWMEELICNTTVERCQELYRKYEMQYAPQPPPTVCQF
ITATEMTLQRYAADINSRLIRQVERLKQEAVTVPVCEDQLKEIERCIKVFLHENGEEGSLSLASVIIS
ALCTLTRRNLMMEGAASSAGEQLVDLTSRDGAWFLEELCSMSGNVTCLVQLLKQCHLVPQDLDIPNPM
EASETVHLANGVYTSLQELNSNFRQIIFPEALRCLMKGEYTLESMLHELDGLIEQTTDGVPLQTLVES
LQAYLRNAAMGLEEETHAHYIDVARLLHAQYGELIQPRNGSVDETPKMSAGQMLLVAFDGMFAQVETA
FSLLVEKLNKMEIPIAWRKIDIIREARSTQVNFFDDDNHRQVLEEIFFLKRLQTIKEFFRLCGTFSKT
LSGSSSLEDQNTVNGPVQIVNVKTLFRNSCFSEDQMAKPIKAFTADFVRQLLIGLPNQALGLTLCSFI
SALGVDIIAQVEAKDFGAESKVSVDDLCKKAVEHNIQIGKFSQLVMNRATVLASSYDTAWKKHDLVRR
LETSISSCKTSLQRVQLHIAMFQWQHEDLLINRPQAMSVTPPPRSAILTSMKKKLHTLSQIETSIATV
QEKLAALESSIEQRLKWAGGANPALAPVLQDFEATIAERRNLVLKESQRASQVTFLCSNIIHFESLRT
RTAEALNLDAALFELIKRCQQMCSFASQFNSSVSELELRLLQRVDTGLEHPIGSSEWLLSAHKQLTQD
MSTQRAIQTEKEQQIETVCETIQNLVDNIKTVLTGHNRQLGDVKHLLKAMAKDEEAALADGEDVPYEN
SVRQFLGEYKSWQDNIQTVLFTLVQAMGQVRSQEHVEMLQEITPTLKELKTQSQSIYNNLVSFASPLV
TDATNECSSPTSSATYQPSFAAAVRSNTGQKTQPDVMSQNARKLIQKNLATSADTPPSTVPGTGKSVA
CSPKKAVRDPKTGKAVQERNSYAVSVWKRVKAKLEGRDVDPNRRMSVAEQVDYVIKEATNLDNLAQLY
EGWTAWV